

Molecular modelling of reaction kinetics

Winter School for Theoretical Chemistry and Spectroscopy
Han-sur-Lesse, 7-11 December 2015

- The Liouville Operator
- MD Integrators

- Rare event simulations
- Reactive flux method

- Free energy methods
- path-metadynamics

Equations of motion

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\Delta t^2}{2m} \mathbf{f}(t) + \frac{\Delta t^3}{3!} \frac{d^3 \mathbf{r}(t)}{dt^3} + \mathcal{O}(\Delta t^4)$$

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \frac{\Delta t^2}{2m} \mathbf{f}(t) - \frac{\Delta t^3}{3!} \frac{d^3 \mathbf{r}(t)}{dt^3} + \mathcal{O}(\Delta t^4)$$

$$\mathbf{r}(t + \Delta t) + \mathbf{r}(t - \Delta t) = 2\mathbf{r}(t) + \frac{\Delta t^2}{m} \mathbf{f}(t) + \mathcal{O}(\Delta t^4)$$

Verlet algorithm

$$\mathbf{r}(t + \Delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \frac{\Delta t^2}{m} \mathbf{f}(t)$$

Velocity Verlet algorithm

$$\mathbf{r}(t + \Delta t) \approx \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\Delta t^2}{2m} \mathbf{f}(t)$$

$$\mathbf{v}(t + \Delta t) \approx \mathbf{v}(t) + \frac{\Delta t}{2m} [\mathbf{f}(t + \Delta t) + \mathbf{f}(t)]$$

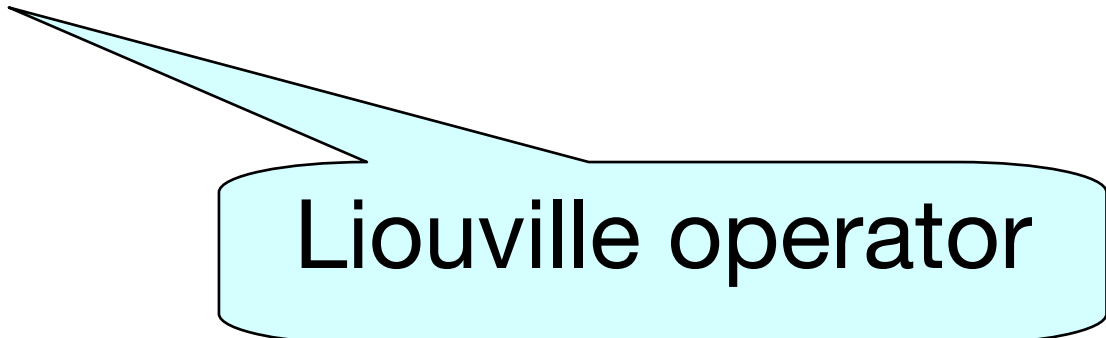
Liouville formulation

$$f(\mathbf{r}^N, \mathbf{p}^N)$$

$$\dot{f} = \dot{\mathbf{r}} \frac{\partial f}{\partial \mathbf{r}} + \dot{\mathbf{p}} \frac{\partial f}{\partial \mathbf{p}}$$

$$iL \equiv \dot{\mathbf{r}} \frac{\partial}{\partial \mathbf{r}} + \dot{\mathbf{p}} \frac{\partial}{\partial \mathbf{p}}$$

$$\frac{df}{dt} = iL f$$



Liouville operator

Solution

$$f(t) = \exp(iLt) f(0)$$

$$i\mathcal{L} \equiv i\mathcal{L}_r + i\mathcal{L}_p = \dot{\mathbf{r}} \frac{\partial}{\partial \mathbf{r}} + \dot{\mathbf{p}} \frac{\partial}{\partial \mathbf{p}}$$

$$f(t) = \exp(i\mathcal{L}_r t) f(0)$$

$$= \exp \left(\dot{\mathbf{r}}(0) t \frac{\partial}{\partial \mathbf{r}} f(0) \right)$$

$$= \sum_{n=0}^{\infty} \frac{(\dot{\mathbf{r}}(0) t)^n}{n!} \frac{\partial^n}{\partial \mathbf{r}^n} f(0)$$

$$= f \left(\mathbf{p}^N(0), (\mathbf{r}(0) + \dot{\mathbf{r}}(\mathbf{0})t) \right)$$

Shift of coordinates

$$\mathbf{r}(0) \rightarrow \mathbf{r}(0) + \dot{\mathbf{r}}(\mathbf{0})t$$

$$f(t) = \exp(i\mathcal{L}_p t) f(0)$$

$$= \exp \left(\dot{\mathbf{p}}(0) t \frac{\partial}{\partial \mathbf{p}} f(0) \right)$$

$$= \sum_{n=0}^{\infty} \frac{(\dot{\mathbf{p}}(0) t)^n}{n!} \frac{\partial^n}{\partial \mathbf{p}^n} f(0)$$

$$= f \left((\mathbf{p}(0) + \dot{\mathbf{p}}(\mathbf{0})t)^N, \mathbf{r}^N(0) \right)$$

Shift of momenta

$$\mathbf{p}(0) \rightarrow \mathbf{p}(0) + \dot{\mathbf{p}}(\mathbf{0})t$$

Taylor expansion

$$i\mathcal{L}_r \Rightarrow \mathbf{r}(0) \rightarrow \mathbf{r}(0) + \dot{\mathbf{r}}(\mathbf{0})t$$

$$i\mathcal{L}_p \Rightarrow \mathbf{p}(0) \rightarrow \mathbf{p}(0) + \dot{\mathbf{p}}(\mathbf{0})t$$

We have *noncommuting* operators!

$$e^{A+B} \neq e^A e^B$$

Trotter identity

$$e^{A+B} = \lim_{P \rightarrow \infty} (e^{A/2P} e^{B/P} e^{A/2P})^P$$

$$e^{A+B} \approx (e^{A/2P} e^{B/P} e^{A/2P})^P$$

$$= \left(e^{i \frac{A}{P}} e^{i \frac{B}{P}} e^{i \frac{A}{P}} \right)^P \quad \frac{A}{P} = \frac{i\mathcal{L}_p t}{P} \quad \frac{B}{P} = \frac{i\mathcal{L}_r t}{P} \quad \Delta t = \frac{t}{P}$$

$$i\mathcal{L}_r \Delta t \Rightarrow \mathbf{r} \rightarrow \mathbf{r} + \dot{\mathbf{r}} \Delta t \quad \left(e^{(i\mathcal{L}_p \Delta t/2)} e^{(i\mathcal{L}_r \Delta t)} e^{(i\mathcal{L}_p \Delta t/2)} \right)^P$$

$$i\mathcal{L}_p \Delta t \Rightarrow \mathbf{p} \rightarrow \mathbf{p} + \dot{\mathbf{p}} \Delta t$$

$$e^{(i\mathcal{L}_p \Delta t/2)} f(\mathbf{p}^N(0), \mathbf{r}^N(0)) = f\left(\left[\mathbf{p}(0) + \frac{\Delta t}{2} \dot{\mathbf{p}}(0)\right]^N, \mathbf{r}^N(0)\right)$$

Velocity Verlet!

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t) \Delta t + \frac{1}{2m} \mathbf{F}(t) \Delta t^2$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \frac{\Delta t}{2m} [\mathbf{F}(t) + \mathbf{F}(t + \Delta t)]$$

$$\mathbf{p}(0) \rightarrow \mathbf{p}(0) + \frac{\Delta t}{2} [\dot{\mathbf{p}}(0) + \dot{\mathbf{p}}(\Delta t)]$$

$$\mathbf{r}(0) \rightarrow \mathbf{r}(0) + \Delta t \dot{\mathbf{r}}\left(\frac{\Delta t}{2}\right) = \mathbf{r}(0) + \Delta t \dot{\mathbf{r}}(0) + \frac{\Delta t^2}{2m} \mathbf{F}(\mathbf{0})$$

Velocity Verlet:

$$e^{(i\mathcal{L}_p\Delta t/2)} e^{(i\mathcal{L}_r\Delta t)} e^{(i\mathcal{L}_p\Delta t/2)}$$

Call force(**fx**)

Do while (**t**<**tmax**)

$$e^{(i\mathcal{L}_p\Delta t/2)} : \mathbf{v} \left(t + \frac{\Delta t}{2} \right) \rightarrow \mathbf{v}(t) + \frac{\Delta t}{2m} \mathbf{f}(\mathbf{t})$$

vx=vx+delt*t*fx/2

$$e^{(i\mathcal{L}_r\Delta t)} : \mathbf{r}(t + \Delta t) \rightarrow \mathbf{r}(t) + \Delta t \mathbf{v}(t + \Delta t/2)$$

x=x+delt*v

Call force(**fx**)

$$e^{(i\mathcal{L}_p\Delta t/2)} : \mathbf{v}(t + \Delta t) \rightarrow \mathbf{v}(t + \Delta t/2) + \frac{\Delta t}{2m} \mathbf{f}(\mathbf{t} + \Delta \mathbf{t})$$

vx=vx+delt*t*fx/2

enddo

Liouville Formulation

Velocity Verlet algorithm:

$$\mathbf{p}(t + \Delta t) = \mathbf{p}(t) + \frac{\Delta t}{2} [\dot{\mathbf{p}}(t) + \dot{\mathbf{p}}(t + \Delta t)]$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \dot{\mathbf{r}}(t) + \frac{\Delta t^2}{2m} \mathbf{F}(t)$$

Three subsequent coordinate transformations in either \mathbf{r} or \mathbf{p} of which the *Jacobian* is one: *Area preserving*

$$\mathbf{p}(t + \Delta t/2) = \mathbf{p}(t) + \frac{\Delta t}{2} \mathbf{F}(\mathbf{r})$$

$$\mathbf{r}(t) = \mathbf{r}(t)$$

$$J_1 = \det \begin{vmatrix} 1 & \frac{\Delta t}{2} \frac{\partial \mathbf{F}(\mathbf{r})}{\partial \mathbf{r}} \\ 0 & 1 \end{vmatrix} = 1$$

$$\mathbf{p}(t + \Delta t/2) = \mathbf{p}(t + \Delta t/2)$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \frac{\Delta t}{m} \mathbf{p}(t + \Delta t/2)$$

$$J_2 = \det \begin{vmatrix} 1 & 0 \\ \Delta t/m & 1 \end{vmatrix} = 1$$

$$\mathbf{p}(t + \Delta t) = \mathbf{p}(t + \Delta t/2) + \frac{\Delta t}{2} \mathbf{F}(\mathbf{r}(t))$$

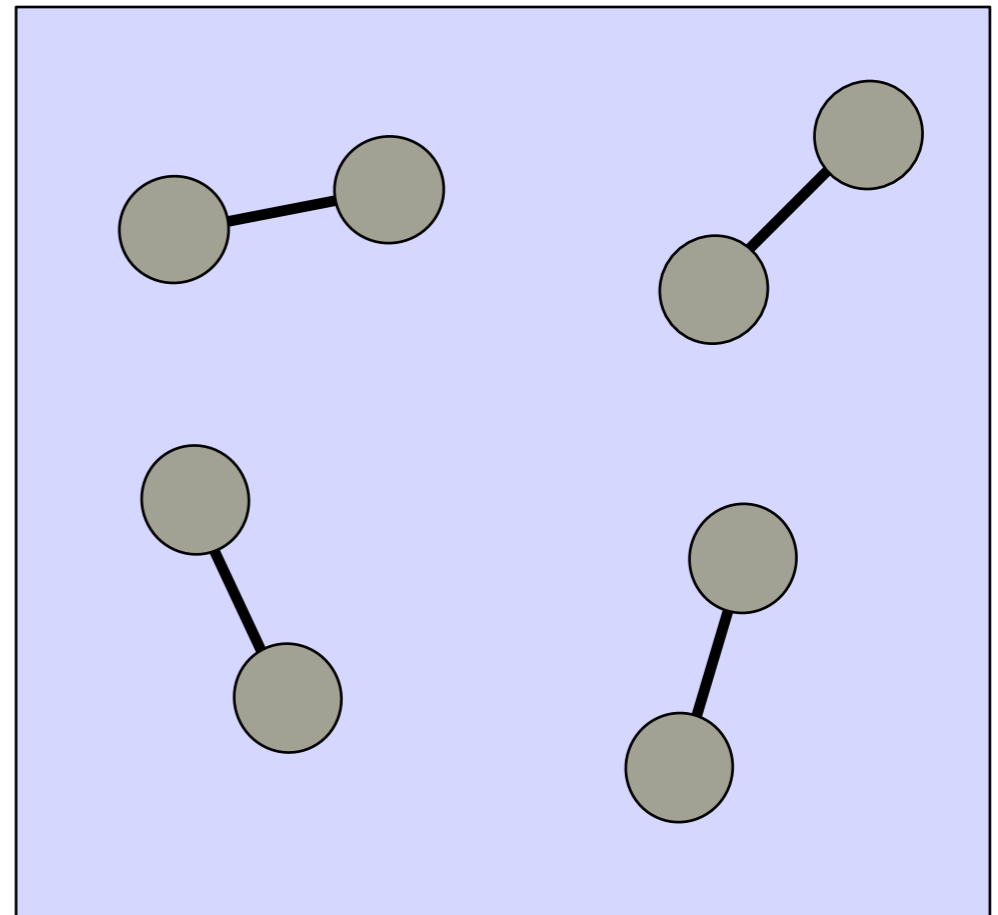
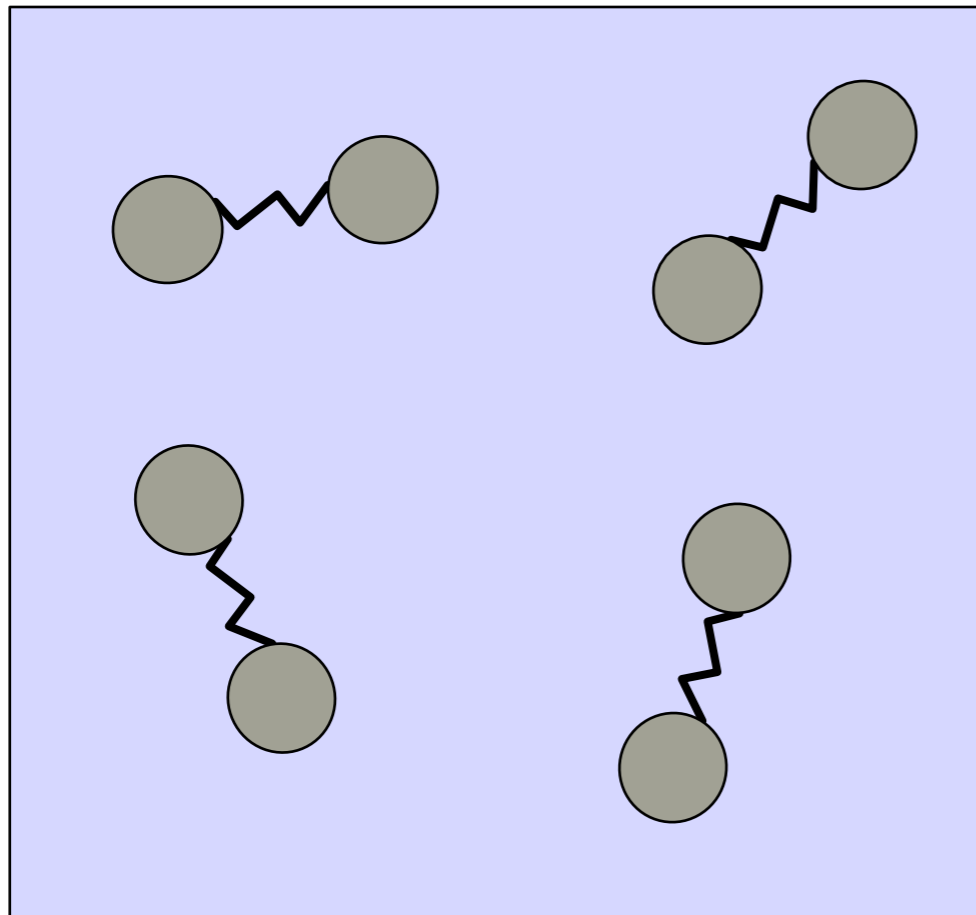
$$\mathbf{r}(t) = \mathbf{r}(t)$$

$$J_3 = \det \begin{vmatrix} 1 & \frac{\Delta t}{2} \frac{\partial \mathbf{F}(\mathbf{r})}{\partial \mathbf{r}} \\ 0 & 1 \end{vmatrix} = 1$$

Other Trotter decompositions are possible!

Multiple time steps

- What to use for stiff potentials:



- Very small time step
- Fixed bond-length: constraints (SHAKE/RATTLE/ROLL)
- reversible reference system propagation algorithm (r-RESPA)

Multiple time steps

$$\mathbf{F} = \mathbf{F}_{\text{short}} + \mathbf{F}_{\text{long}}$$

$$i\mathcal{L} \equiv i\mathcal{L}_r + i\mathcal{L}_p = \mathbf{v} \frac{\partial}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m} \frac{\partial}{\partial \mathbf{v}}$$

$$i\mathcal{L} \equiv i\mathcal{L}_{\text{short}} + i\mathcal{L}_{\text{long}}$$

$$i\mathcal{L}_{\text{short}} = \frac{\mathbf{F}_{\text{short}}}{m} \frac{\partial}{\partial \mathbf{v}} \quad i\mathcal{L}_{\text{long}} = \frac{\mathbf{F}_{\text{long}}}{m} \frac{\partial}{\partial \mathbf{v}}$$

Trotter expansion:

$$e^{i(\mathcal{L}_{\text{long}} + \mathcal{L}_{\text{short}} + \mathcal{L}_r)\Delta t} \approx e^{i\mathcal{L}_{\text{long}}\Delta t/2} e^{i(\mathcal{L}_{\text{short}} + \mathcal{L}_r)\Delta t} e^{i\mathcal{L}_{\text{long}}\Delta t/2}$$

Introduce: $\delta t = \Delta t/n$

$$\approx e^{i\mathcal{L}_{\text{long}}\Delta t/2} \left[e^{i\mathcal{L}_{\text{short}}\delta t/2} e^{i\mathcal{L}_r\delta t} e^{i\mathcal{L}_{\text{short}}\delta t/2} \right]^n e^{i\mathcal{L}_{\text{long}}\Delta t/2}$$

$$e^{i(\mathcal{L}_{\text{long}} + \mathcal{L}_{\text{short}} + \mathcal{L}_r)\Delta t} \approx e^{i\mathcal{L}_{\text{long}}\Delta t/2} \left[e^{i\mathcal{L}_{\text{short}}\delta t/2} e^{i\mathcal{L}_r\delta t} e^{i\mathcal{L}_{\text{short}}\delta t/2} \right]^n e^{i\mathcal{L}_{\text{long}}\Delta t/2}$$

$$i\mathcal{L}_{\text{long}}\Delta t/2 \Rightarrow v \rightarrow v + F_{\text{long}}\Delta t/2m$$

$$i\mathcal{L}_{\text{short}}\delta t/2 \Rightarrow v \rightarrow v + F_{\text{short}}\delta t/2m$$

$$i\mathcal{L}_r\delta t \Rightarrow r \rightarrow r + v\delta t$$

First

$$e^{(i\mathcal{L}_{\text{long}}\Delta t/2)} f[r(0), v(0)] = f[r(0), v(0) + F_{\text{long}}(0)\Delta t/2m]$$

Now n times:

$$\left[e^{i\mathcal{L}_{\text{short}}\delta t/2} e^{i\mathcal{L}_r\delta t} e^{i\mathcal{L}_{\text{short}}\delta t/2} \right]^n f[r(0), v(0) + F_{\text{long}}(0)\Delta t/2m]$$

$$e^{(i\mathcal{L}_{long}\Delta t/2)} : \mathbf{v} \left(t + \frac{\Delta t}{2} \right) \rightarrow \mathbf{v}(t) + \frac{\Delta t}{2m} \mathbf{f}_{long}(t)$$

Call force(fx_long, f_short)

vx=vx+delt*fx_long/2

Do ddt=1,n

$$e^{(i\mathcal{L}_{short}\delta t/2)} : \mathbf{v} \left(t + \frac{\Delta t}{2} + \frac{\delta t}{2} \right) \rightarrow \mathbf{v} \left(t + \frac{\Delta t}{2} \right) + \frac{\delta t}{2m} \mathbf{f}_{short}(t)$$

vx=vx+ddelt*fx_short/2

$$e^{(i\mathcal{L}_r\Delta t)} : \mathbf{r} (t + \delta t) \rightarrow \mathbf{r}(t) + \delta t \mathbf{v} (t + \Delta t/2 + \delta t/2)$$

x=x+ddelt*vx

Call force_short(fx_short)

$$e^{(i\mathcal{L}_{short}\delta t/2)} : \mathbf{v} \left(t + e^{i\mathcal{L}_{long}\Delta t/2} \left[e^{i\mathcal{L}_{short}\delta t/2} e^{i\mathcal{L}_r\delta t} e^{i\mathcal{L}_{short}\delta t/2} \right]^n e^{i\mathcal{L}_{long}\Delta t/2} \right)$$

vx=vx+ddelt*fx

enddo

$$e^{(i\mathcal{L}_{long}\Delta t/2)} : \mathbf{v} (t + \Delta t) \rightarrow \mathbf{v} (t + \frac{\Delta t}{2}) + \frac{\Delta t}{2m} \mathbf{f}_{long}(t + \frac{\Delta t}{2})$$

$$e^{i\mathcal{L}_{long}\Delta t/2} \left[e^{i\mathcal{L}_{short}\delta t/2} e^{i\mathcal{L}_r\delta t} e^{i\mathcal{L}_{short}\delta t/2} \right]^n e^{i\mathcal{L}_{long}\Delta t/2}$$

$$i\mathcal{L}_{long}\Delta t/2 \Rightarrow v \rightarrow v + F_{long}\Delta t/2m$$

$$i\mathcal{L}_{short}\delta t/2 \Rightarrow v \rightarrow v + F_{short}\delta t/2m$$

$$i\mathcal{L}_r\delta t \Rightarrow r \rightarrow r + v\delta t$$

Algorithm 29 (Multiple Time Step)

```
subroutine
+   multi(f_long, f_short)

vx=vx+0.5*delt*f_long
do it=1,n
    vx=vx+0.5*(delt/n)*f_short
    x=x+(delt/n)*vx
    call force_short(f_short)
    vx=vx+0.5*(delt/n)*f_short
enddo
call force_all(f_long, f_short)
vx=vx+0.5*delt*f_long
return
end
```

Multiple time step, `f_long` is the long-range part and `f_short` the short-range part of the force velocity Verlet with time step Δt loop for the small time step velocity Verlet with timestep $\Delta t/n$

short-range forces

all forces

Summary part 1/3

- Liouville formulation
- Velocity Verlet Algorithm
- Area preserving, time-reversible, energy conserving,...
- Multi-timestep algorithm

- The Liouville Operator
- MD Integrators

- Rare event simulations
- Reactive flux method

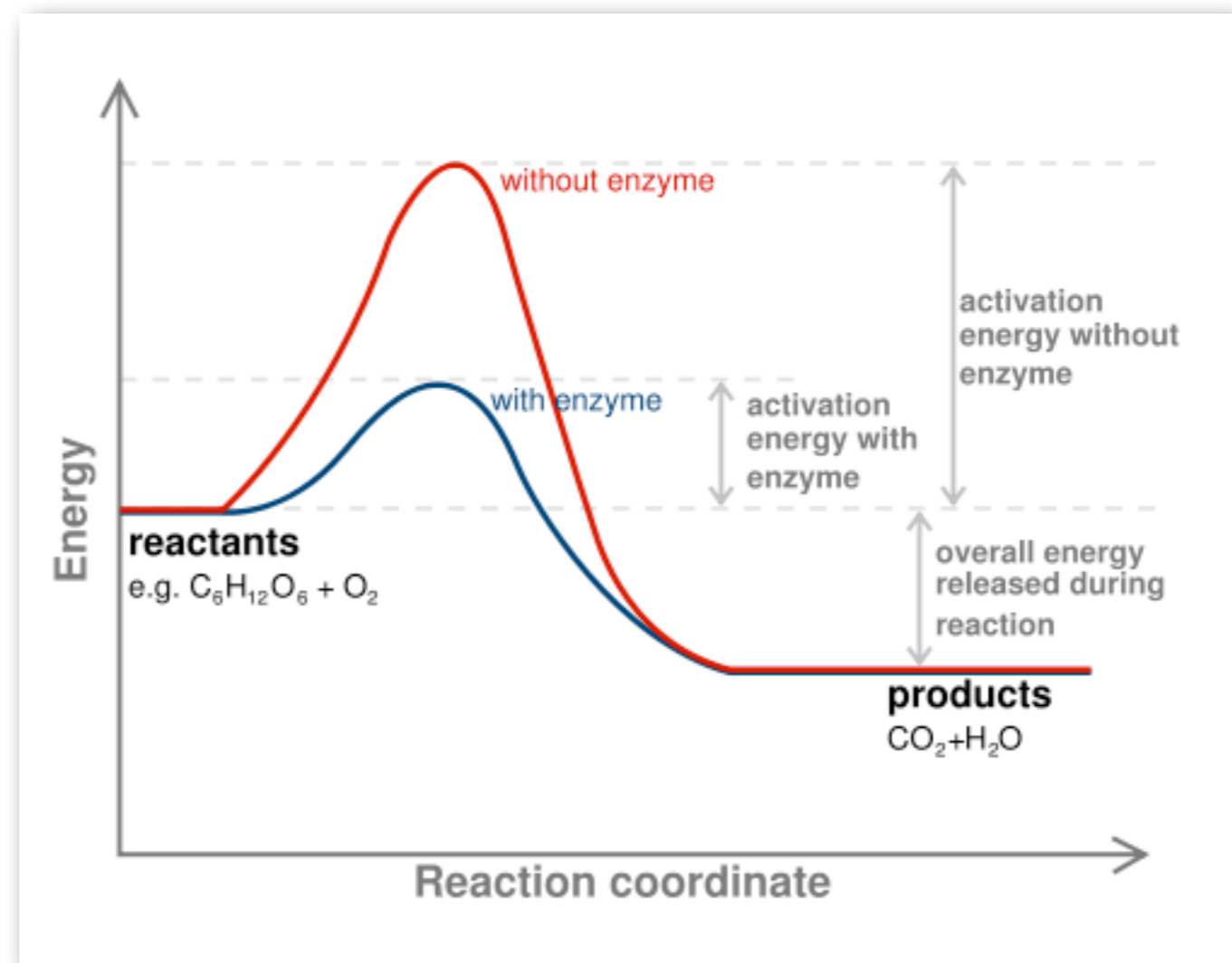
- Free energy methods
- path-metadynamics

Rare events

In real life (experiment)

- no direct access to free energy
- reaction rate measurement

$$k = k_0 e^{-\Delta G/kT}$$



Rare event simulation

Macroscopic phenomenological theory

Chemical reaction: $A \leftrightarrow B$

$$\frac{dc_A(t)}{dt} = -k_{A \rightarrow B}c_A(t) + k_{B \rightarrow A}c_B(t)$$

$$\frac{dc_B(t)}{dt} = +k_{A \rightarrow B}c_A(t) - k_{B \rightarrow A}c_B(t)$$

Total number of molecules:

$$\frac{d[c_A(t) + c_B(t)]}{dt} = 0$$

Equilibrium:

$$\frac{dc_A(t)}{dt} = \frac{dc_B(t)}{dt} = 0 \quad \frac{\langle c_A \rangle}{\langle c_B \rangle} = \frac{k_{B \rightarrow A}}{k_{A \rightarrow B}}$$

Rare event simulation

Macroscopic phenomenological theory

Make a small perturbation:

$$c_A(t) = \langle c_A \rangle + \Delta c_A(t) \quad c_B(t) = \langle c_B \rangle + \Delta c_A(t)$$

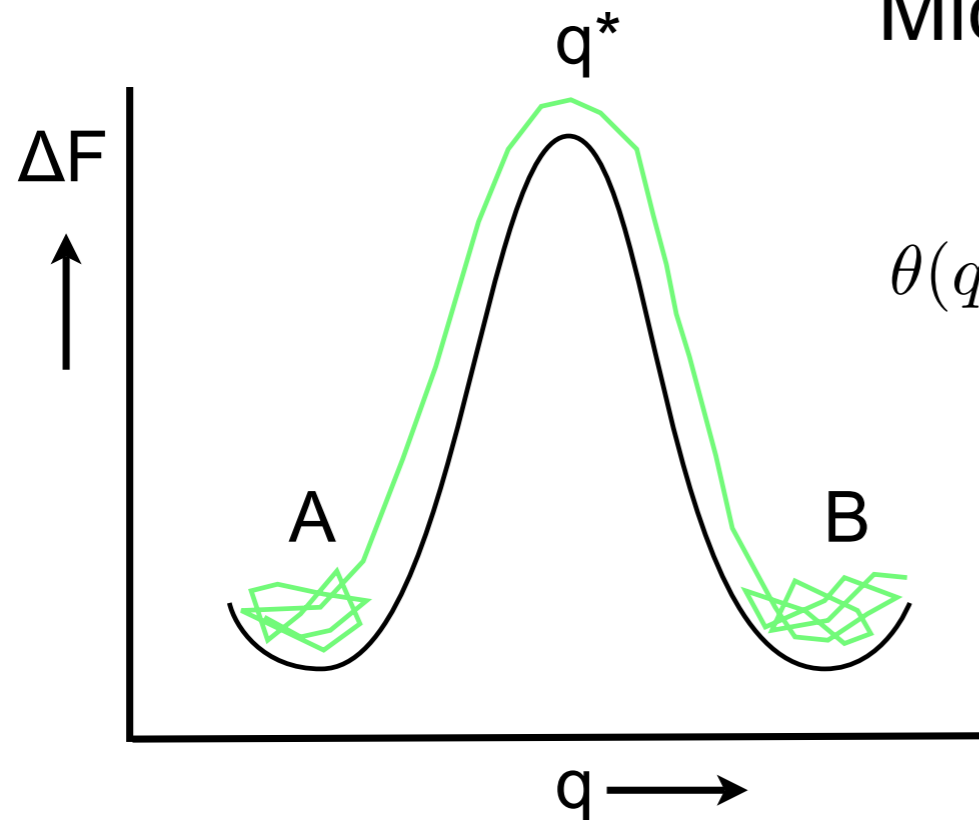
$$\frac{d\Delta c_A(t)}{dt} = -k_{A \rightarrow B} \Delta c_A(t) - k_{B \rightarrow A} \Delta c_A(t)$$

$$\begin{aligned} \Delta c_A(t) &= \Delta c_A(0) \exp[-(k_{A \rightarrow B} + k_{B \rightarrow A})t] \\ &= \Delta c_A(0) \exp[-t/\tau] \end{aligned}$$

$$\begin{aligned} \tau &= (k_{A \rightarrow B} + k_{B \rightarrow A})^{-1} \\ &= k_{A \rightarrow B}^{-1} (1 + \langle c_A \rangle / \langle c_B \rangle)^{-1} = \frac{\langle c_B \rangle}{k_{A \rightarrow B}} \end{aligned}$$

Rare event simulation

Microscopic linear response theory



$$\theta(q - q^*) = \begin{cases} 0 & \text{if } q - q^* < 0 \quad (\text{Reactant A}) \\ 1 & \text{if } q - q^* > 0 \quad (\text{Product B}) \end{cases}$$

Perturbation

add bias to increase concentration c_A

$$H = H_0 - \epsilon g_A(q - q^*)$$

$$g_A(q - q^*) = 1 - \theta(q - q^*) = \theta(q^* - q)$$

$$\Delta c_A = \langle c_A \rangle_\epsilon - \langle c_A \rangle_0$$

$$\Delta c_A = \langle g_A \rangle_\epsilon - \langle g_A \rangle_0$$

$\langle g_A \rangle \rightarrow$ probability to be in state A

Linear response theory: static

$$\langle \Delta A \rangle = \langle A \rangle - \langle A \rangle_0 \quad H = H_0 - \epsilon B$$

$$\langle A \rangle_0 = \frac{\int d\Gamma A \exp[-\beta(H_0)]}{\int d\Gamma \exp[-\beta(H_0)]} \quad \langle A \rangle = \frac{\int d\Gamma A \exp[-\beta(H_0 - \epsilon B)]}{\int d\Gamma \exp[-\beta(H_0 - \epsilon B)]}$$

$$\begin{aligned} \left\langle \frac{\partial \Delta A}{\partial \epsilon} \right\rangle &= \frac{\int d\Gamma \beta AB \exp[-\beta(H_0 - \epsilon B)] \int d\Gamma \exp[-\beta(H_0 - \epsilon B)]}{\left\{ \int d\Gamma \exp[-\beta(H_0 - \epsilon B)] \right\}^2} \\ &- \frac{\int d\Gamma A \exp[-\beta(H_0 - \epsilon B)] \int d\Gamma \beta B \exp[-\beta(H_0 - \epsilon B)]}{\left\{ \int d\Gamma \exp[-\beta(H_0 - \epsilon B)] \right\}^2} \\ &= \beta \{ \langle AB \rangle_0 - \langle A \rangle_0 \langle B \rangle_0 \} \end{aligned}$$

Very small perturbation: linear response theory

$$\Delta c_A = \langle g_A \rangle_\epsilon - \langle g_A \rangle_0 \quad H = H_0 - \epsilon g_A(q - q^*)$$

How does the response (Δc) depend on the perturbation ($\Delta \epsilon$)?

$$\begin{aligned} \frac{d\Delta c_A}{d\epsilon} &= \beta \left(\langle (g_A)^2 \rangle_0 - \langle g_A \rangle_0^2 \right) \\ &= \beta \left(\langle g_A \rangle_0 \left(1 - \langle g_A \rangle_0 \right) \right) \\ &= \beta \left(\langle c_A \rangle_0 \left(1 - \langle c_A \rangle_0 \right) \right) = \beta \langle c_A \rangle_0 \langle c_B \rangle_0 \end{aligned}$$

Outside the barrier

$$g_A = 0 \text{ or } 1$$

$$g_A(x)g_A(x) = g_A(x)$$

Switch of the perturbation: dynamic linear response

$$\begin{aligned} \Delta c_A(t) &= \Delta c_A(0) \frac{\Delta g_A(0) \Delta g_A(t)}{\langle c_A \rangle \langle c_B \rangle} \\ &= \Delta c_A(0) \exp[-t/\tau] \end{aligned}$$

holds for sufficiently long times

$$\exp[-t/\tau] = \frac{\langle \Delta g_A(0) \Delta g_A(t) \rangle}{\langle c_A \rangle \langle c_B \rangle}$$

Derivative

$$-\frac{1}{\tau} \exp[-t/\tau] = \frac{\langle g_A(0) \dot{g}_A(t) \rangle}{\langle c_A \rangle \langle c_B \rangle} = \frac{\langle \dot{g}_A(0) g_A(t) \rangle}{\langle c_A \rangle \langle c_B \rangle}$$

Δ has disappeared
because of derivative

For sufficiently short t

$$k_{A \rightarrow B}(t) = \frac{\langle \dot{g}_A(0) g_A(t) \rangle}{\langle c_A \rangle}$$

$$\dot{g}_A(q - q^*) = \dot{q} \frac{\partial g_A(q - q^*)}{\partial q} = -\dot{q} \frac{\partial g_B(q - q^*)}{\partial q}$$

$$k_{a \rightarrow B}(t) = \frac{\left\langle \dot{q}(0) \frac{\partial g_B(q(0) - q^*)}{\partial q} g_B(t) \right\rangle}{\langle c_A \rangle}$$

Stationary

$$\frac{d}{dt} \langle A(t)B(t+t') \rangle = 0$$

$$\langle A(t)\dot{B}(t+t') \rangle + \langle \dot{A}(t)B(t+t') \rangle = 0$$

$$\langle A(t)\dot{B}(t+t') \rangle = -\langle \dot{A}(t)B(t+t') \rangle$$

Eyring's transition state theory

$$\begin{aligned}k_{a \rightarrow B}(t) &= \frac{\left\langle \dot{q}(0) \frac{\partial g_B(q(0) - q^*)}{\partial q} g_B(t) \right\rangle}{\langle c_A \rangle} \\ &= \frac{\langle \dot{q}(0) \delta(q(0) - q^*) \theta(q(t) - q^*) \rangle}{\langle \theta(q^* - q) \rangle}\end{aligned}$$

Correlation between velocity of states that are at the top of the barrier at $t=0$ and in the product state B some time t later.

Let us consider the limit $t \rightarrow 0^+$:


$$\begin{aligned}\lim_{t \rightarrow 0^+} &= \theta(q(t) - q^*) = \theta(\dot{q}(t)) \\ k_{a \rightarrow B}^{\text{TST}}(t) &= \frac{\langle \dot{q}(0) \delta(q(0) - q^*) \theta(\dot{q}) \rangle}{\langle \theta(q^* - q) \rangle}\end{aligned}$$


Bennett-Chandler approach

(or Reactive flux method)

$$k_{a \rightarrow B}(t) = \frac{\langle \dot{q}(0) \delta(q(0) - q^*) \theta(q(t) - q^*) \rangle}{\langle \theta(q^* - q) \rangle}$$

$$k_{a \rightarrow B}(t) = \frac{\langle \dot{q}(0) \delta(q(0) - q^*) \theta(q(t) - q^*) \rangle}{\langle \delta(q(0) - q^*) \rangle} \times \frac{\delta(q(0) - q^*)}{\theta(q^* - q)}$$

 Conditional average: $\dot{q}(0) \theta(q(t) - q^*)$
given that we start on top of barrier

 Probability to find q
on barrier top

Computational scheme:

- Determine the probability with free energy calculation
- Compute conditional average from “shooting” trajectories from barrier top

- The Liouville Operator
- MD Integrators

- Rare event simulations
- Reactive flux method

- Free energy methods
- path-metadynamics

Free energy methods

It is very (very) difficult to compute (or measure) absolute thermodynamic properties, such as free energy and entropy, that depend on the size of the phase space.

But we can compute relative free energies, in particular free energy differences between thermodynamic states.

Reaction equilibrium constants $A \longleftrightarrow B$

$$K = \frac{[B]}{[A]} = \frac{p_B}{p_A} = \exp[-\beta(G_B - G_A)]$$

Examples:

- Chemical reactions, catalysis, isomerization, etc...
- Protein folding, ligand binding affinity, protein-protein association
- Phase diagrams, coexistence lines, critical points, transitions

Free energy perturbation

$$\Delta\beta F = -\ln(Q_B/Q_A) = -\ln\left(\frac{\int ds^N \exp(-\beta U_B)}{\int ds^N \exp(-\beta U_A)}\right)$$

$$\begin{aligned}\Delta F &= -k_B T \ln\left(\frac{\int ds^N \exp(-\beta U_A) \exp(-\beta \Delta U)}{\int ds^N \exp(-\beta U_A)}\right) \\ &= -k_B T \ln \langle \exp(-\beta \Delta U) \rangle_A\end{aligned}$$

$$\Delta F = -k_B T \ln \langle \exp(-\beta \Delta U) \rangle_A = -k_B T \ln \langle \exp(\beta \Delta U) \rangle_B$$

Sampling problems may lead to hysteresis between the two samples

Umbrella Sampling

Bias the sampling along an order parameter q

Add and subtract bias potential $w(q)$:

$$P(q) = \frac{\int d\mathbf{r}^N \delta(q'(\mathbf{r}^N) - q) \exp[-\beta(U(\mathbf{r}^N) + w(q') - w(q))]}{\int d\mathbf{r}^N \exp[-\beta(U(\mathbf{r}^N) + w(q') - w(q))]}$$

$$P(q) = \frac{\int d\mathbf{r}^N \delta(q'(\mathbf{r}^N) - q) \exp[-\beta(U(\mathbf{r}^N) + w(q'))] \exp(\beta w(q))}{\int d\mathbf{r}^N \exp[-\beta(U(\mathbf{r}^N) + w(q'))] \exp(\beta w(q))}$$

$$P(q) = \frac{\langle \delta(q'(\mathbf{r}^N) - q) \exp(\beta w(q')) \rangle_{\text{biased}}}{\langle \exp(\beta w(q')) \rangle_{\text{biased}}}$$

$$P(q) = \frac{\exp(\beta w(q))}{\langle \exp(\beta w(q')) \rangle_{\text{biased}}} P_{\text{biased}}(q)$$

- Let $w(q)$ be a good guess of minus the free energy $F(q)$, or
- Choose $w(q)$ to confine sampling to a specific window along q .
- Or do both.

$$F(q) = k_B T \ln P(q) = -k_B T \ln P_{\text{biased}}(q) - w(q) + \text{const}$$

Constrained MD

The derivative of the free energy $F(\lambda)$ with respect to λ can be written as an ensemble average.

$$\begin{aligned}\left(\frac{\partial F(\lambda)}{\partial \lambda}\right)_{\text{NVT}} &= \frac{\int d\mathbf{r}^N (\partial U(\lambda)/\partial \lambda) \exp[-\beta U(\lambda)]}{\int d\mathbf{r}^N \exp[-\beta U(\lambda)]} \\ &= \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda}\end{aligned}$$

The free energy difference between states A and B can then be obtained by thermodynamic integration

$$F(\lambda_B) - F(\lambda_A) = \int_{\lambda_A}^{\lambda_B} d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda}$$

In the case “hard” constraints are used, additional corrections are needed to unbiased for sampling in a constraint ensemble (instead of the actual NVT ensemble)

Steered MD

Mechanical work to bring the system from state A to state B

$$W_{A \rightarrow B} \geq \Delta F_{A \rightarrow B}$$

Jarzynski's equality

$$\langle \exp[-\beta W_{A \rightarrow B}] \rangle_A = \exp[-\beta \Delta F_{A \rightarrow B}]$$

Surprisingly, we can obtain the equilibrium free energy difference from a non-equilibrium simulation, in which we force the system in a finite time to move from A to B.

Although, this may sound as a *free lunch* method, note that it requires sampling an exponential distribution of the work. For infinitely slow switching from A to B, the system is always in equilibrium so that a single simulation gives ΔF . But the faster the switching the more rare are the important low-work contributions to the average, so that many steered simulations are required for convergence.

Metadynamics

Escaping free-energy minima,
Laio and Parrinello, PNAS (2002)

$$V(t, s) = \sum_{t' < t} H_{t'} \prod_{\alpha} \exp \left[\frac{-(s_{\alpha} - s_{\alpha}^{t'})^2}{2\delta_{\alpha}^2 W^2} \right]$$

- The metadynamics biasing potential “grows” with time, by adding relatively small repulsive Gaussian potentials.
- The potentials placed at visited points in the space of order parameters (collective variables) enhance sampling of unexplored regions.
- The Gaussian “hills” accumulate in the free energy minima, until the counter-balance the basins and allow the system to escape to product states, where the process repeats
- The biasing potential is an estimator of the free energy.

The error depends on the height, width, and time interval of the added Gaussians, and on the diffusion, temperature and order parameter space of the system.

$$\epsilon = C_d \sqrt{\frac{HWS}{D\Delta\tau\beta}}$$

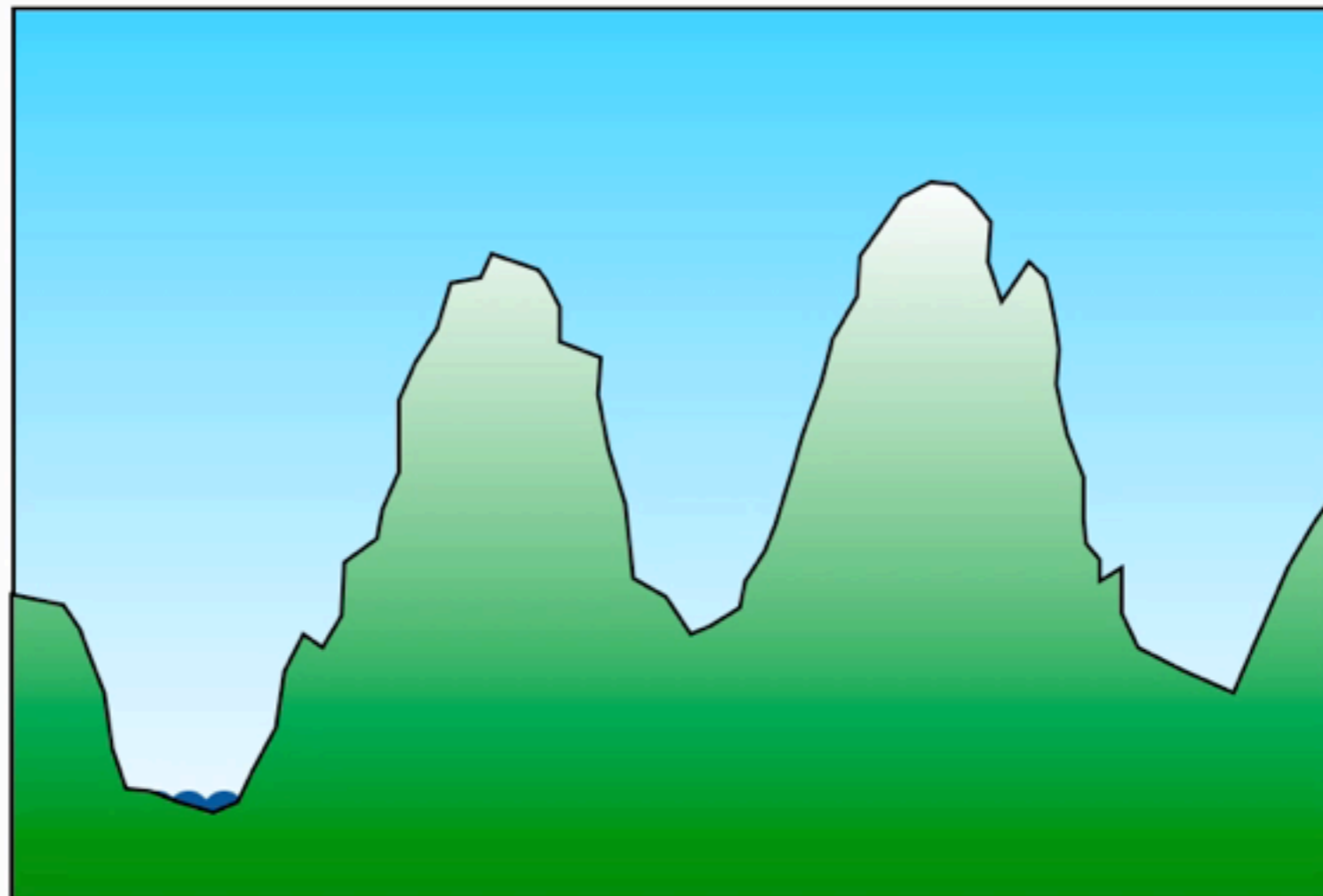
- The Liouville Operator
- MD Integrators

- Rare event simulations
- Reactive flux method

- Free energy methods
- path-metadynamics

A Dutch perspective on escaping free energy minima

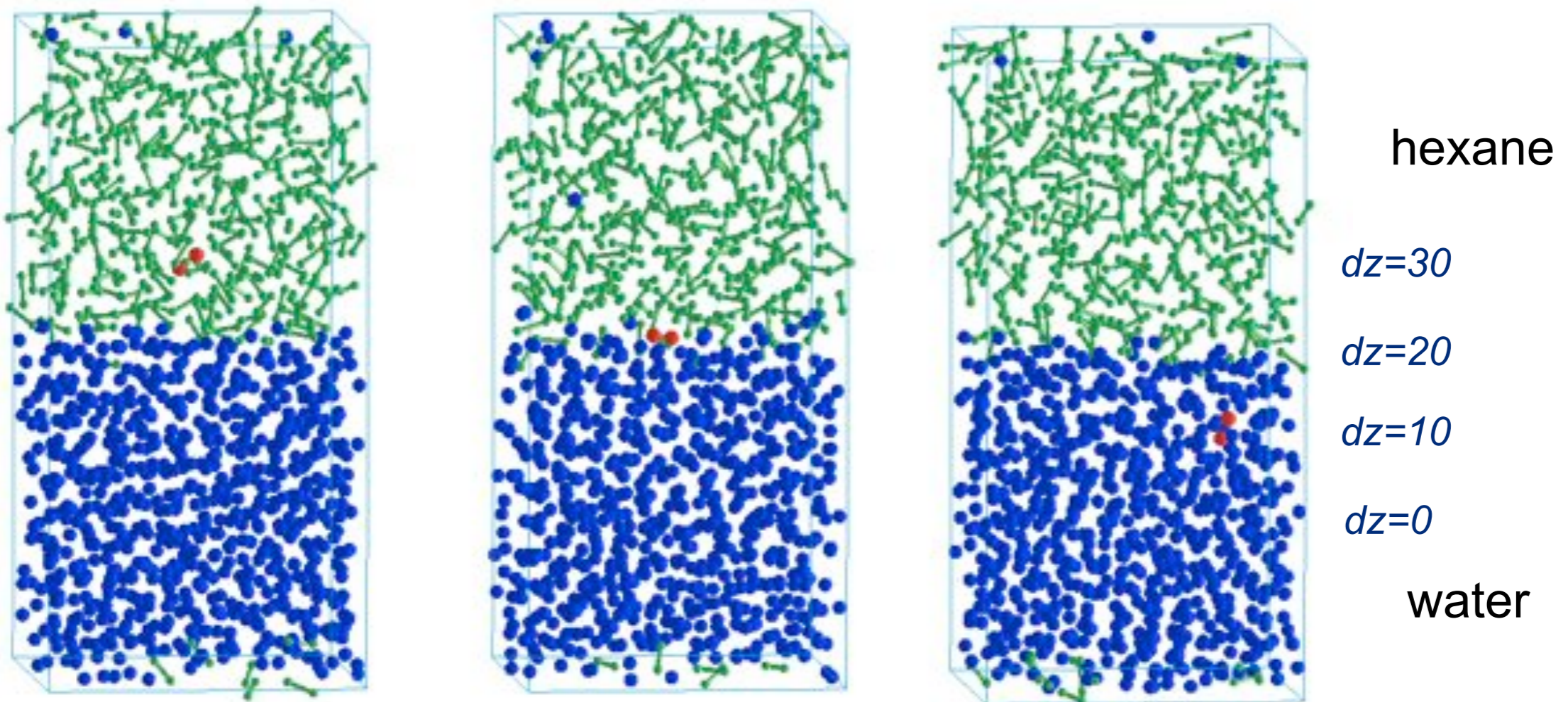
Metadynamics, Laio and Parrinello, PNAS (2002)



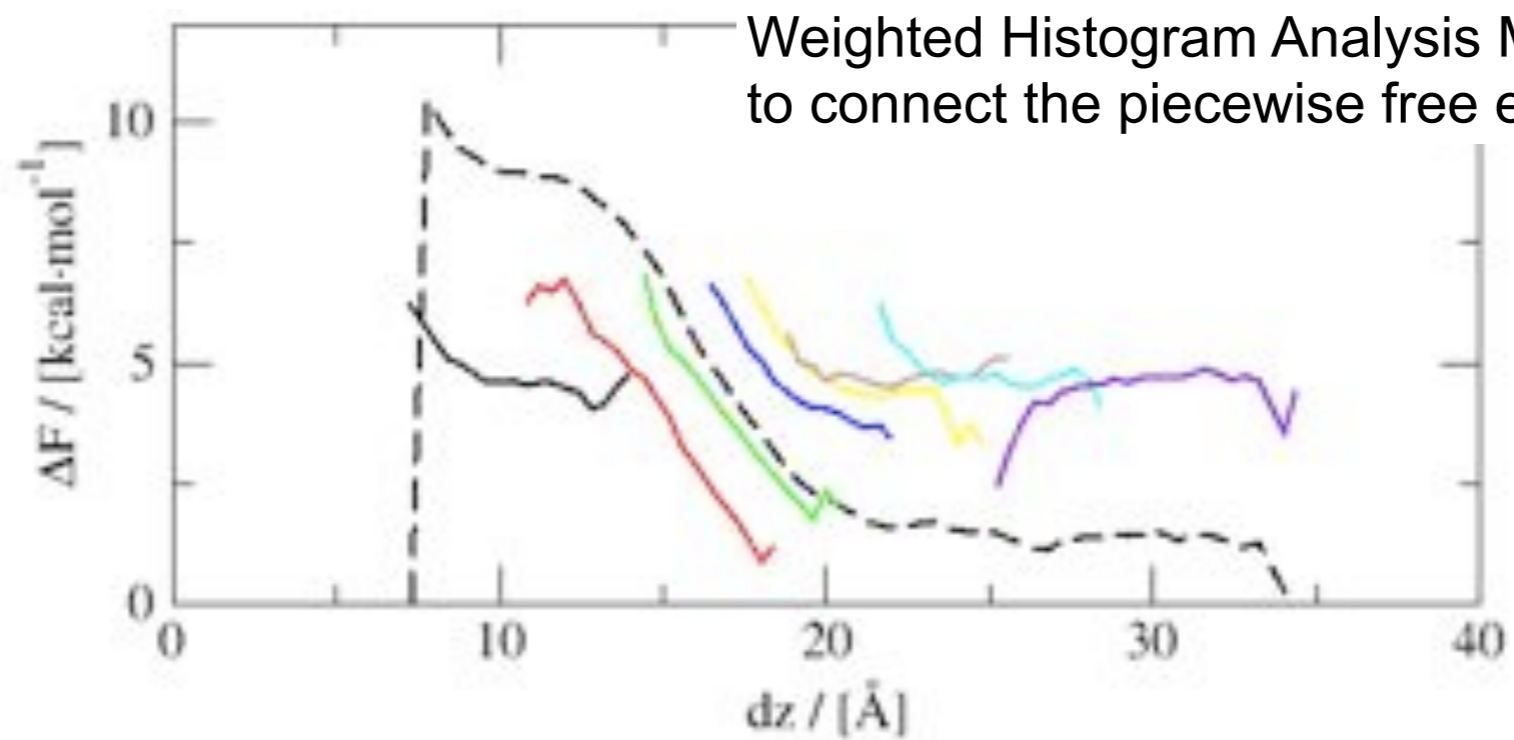
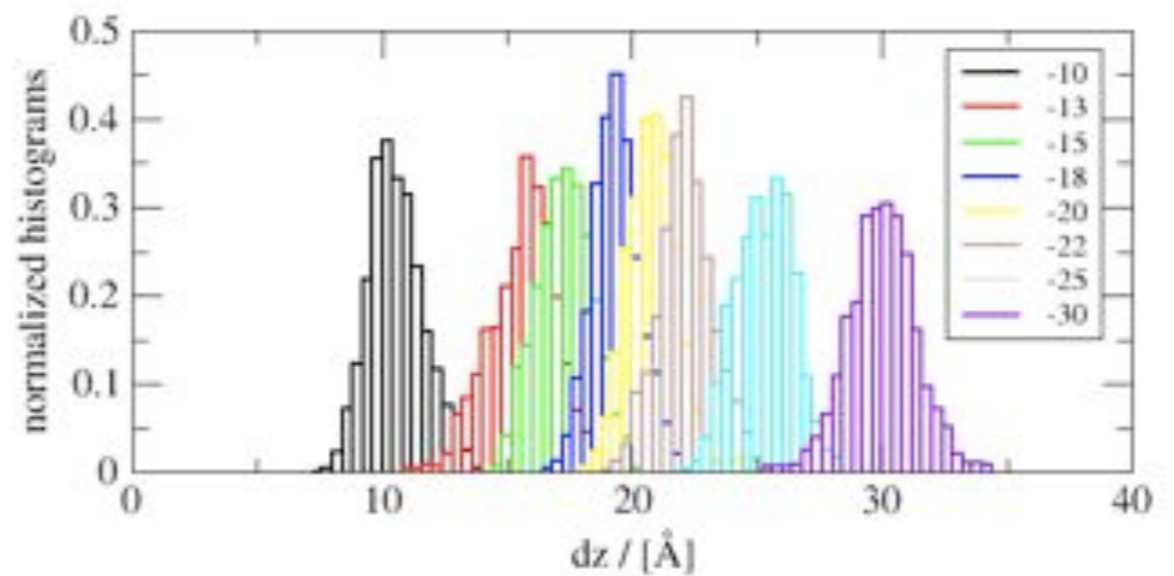
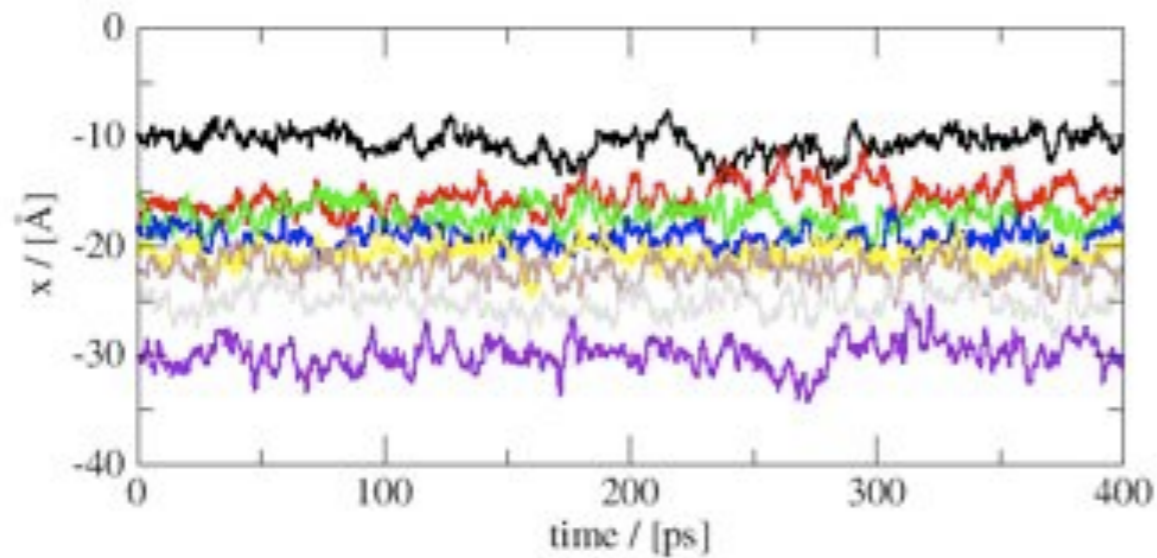
Example

Relative solvation free energy of hexane molecule

The solvation free energy free difference of a solute in different solvents is here used a target property to parameterized a coarse-grain forcefield.



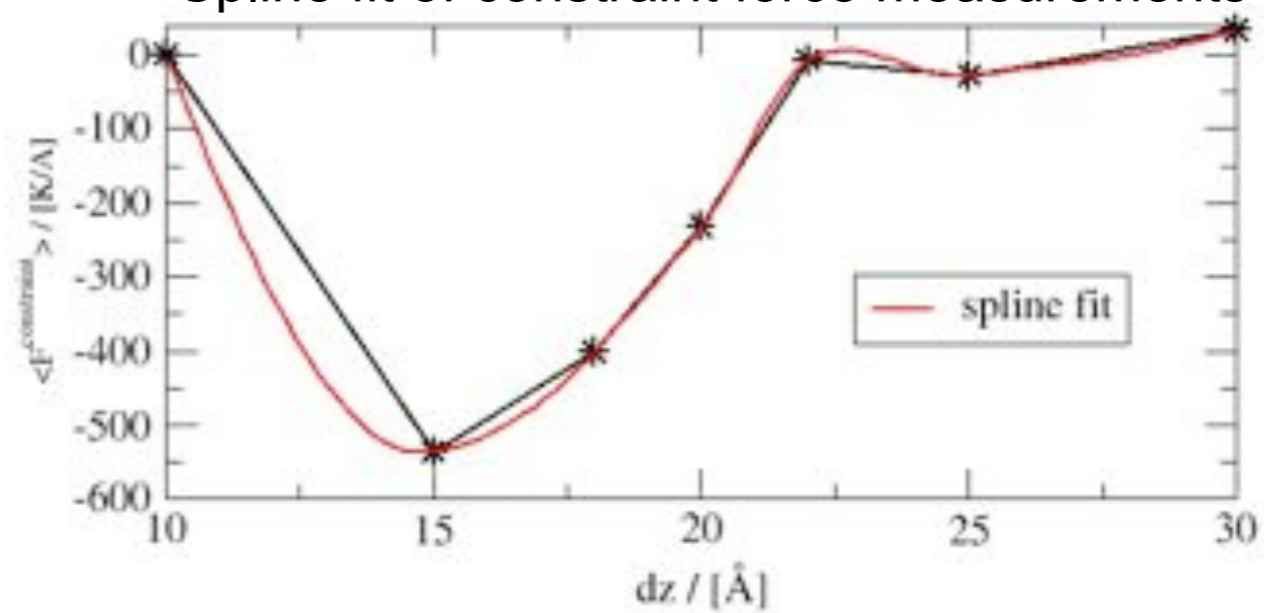
Umbrella Sampling



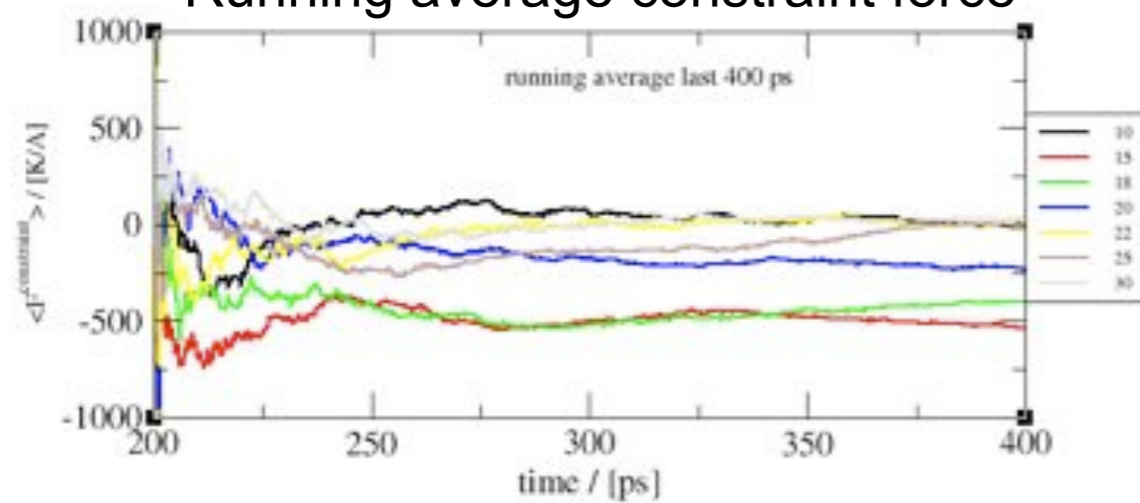
Weighted Histogram Analysis Method (WHAM)
to connect the piecewise free energy curves

Constrained MD

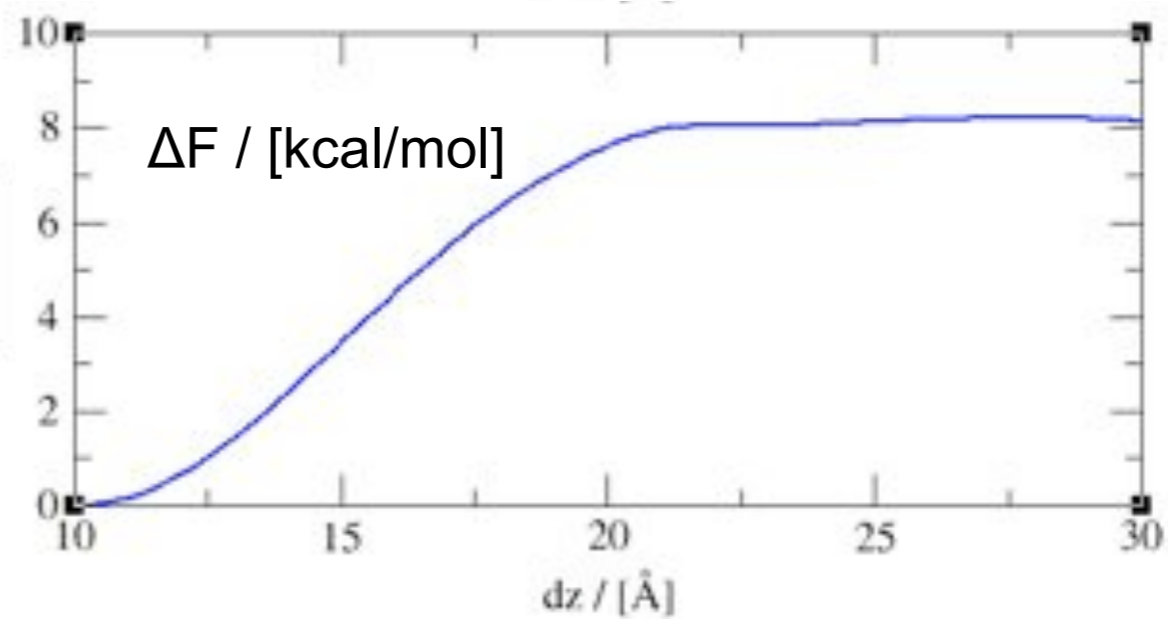
Spline fit of constraint force measurements



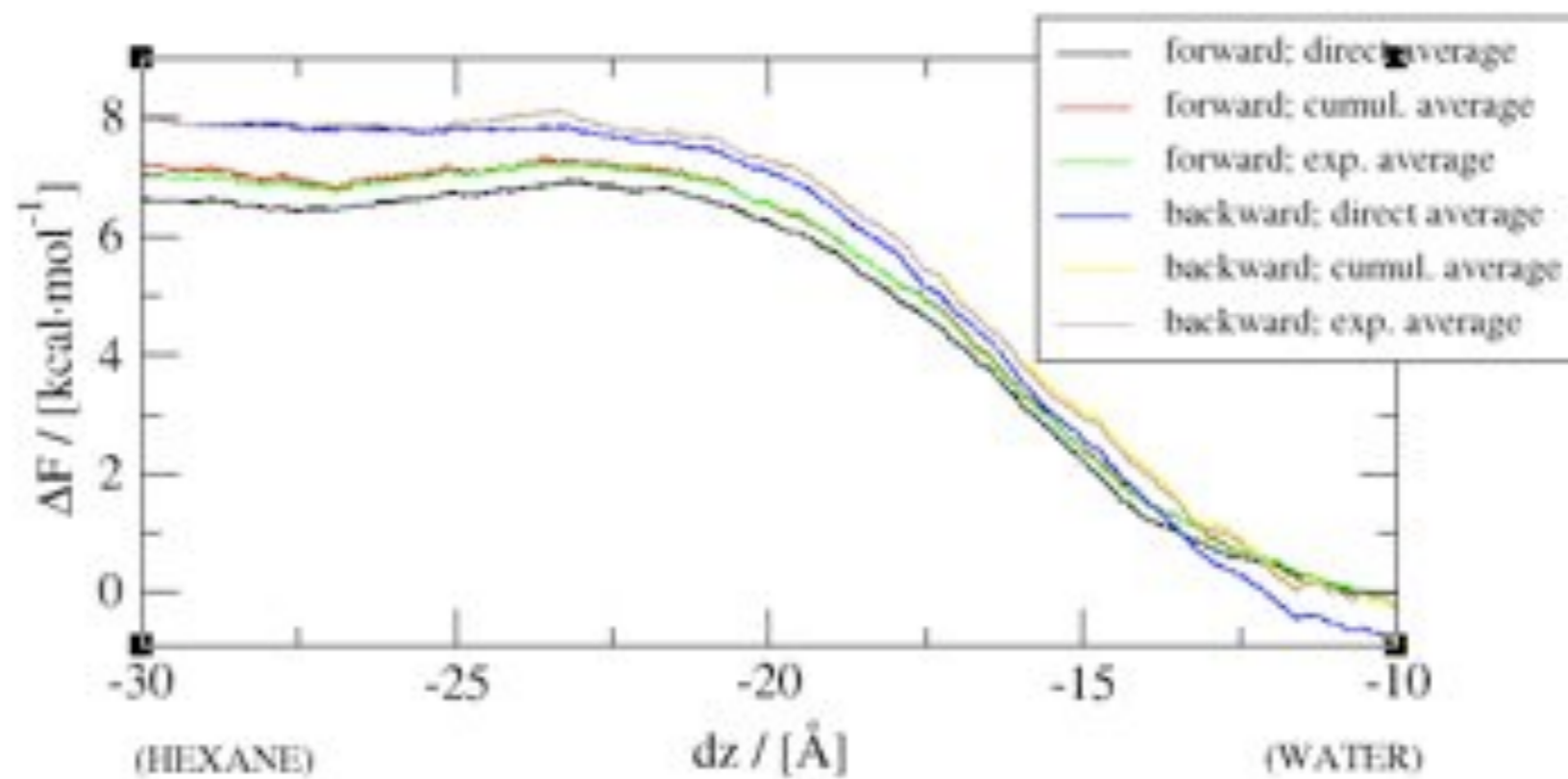
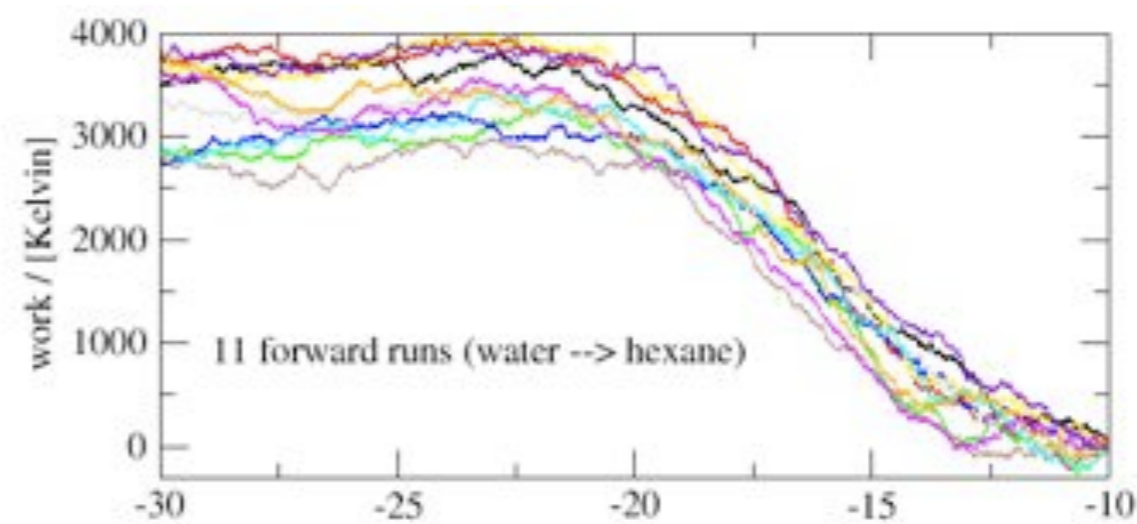
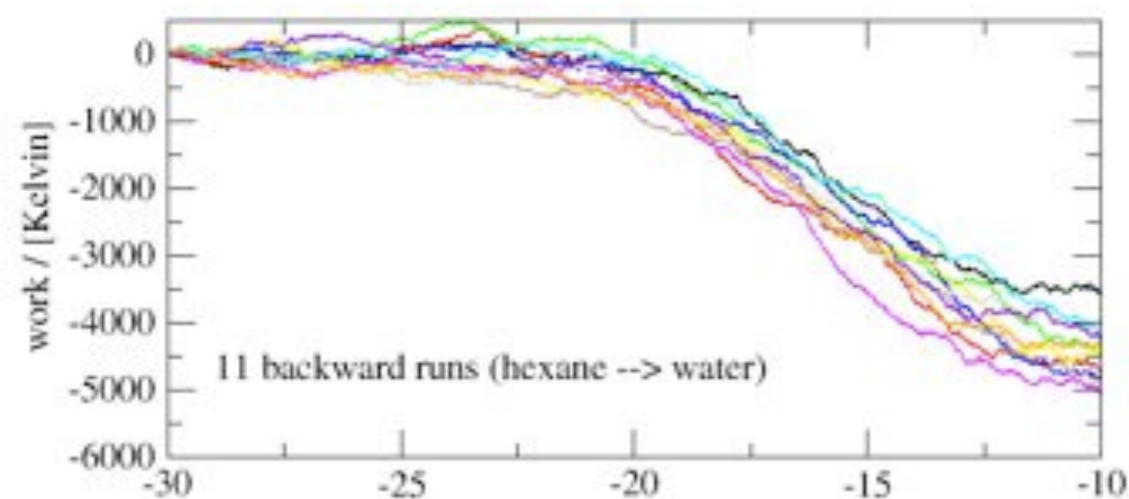
Running average constraint force



$\Delta F / [\text{kcal/mol}]$



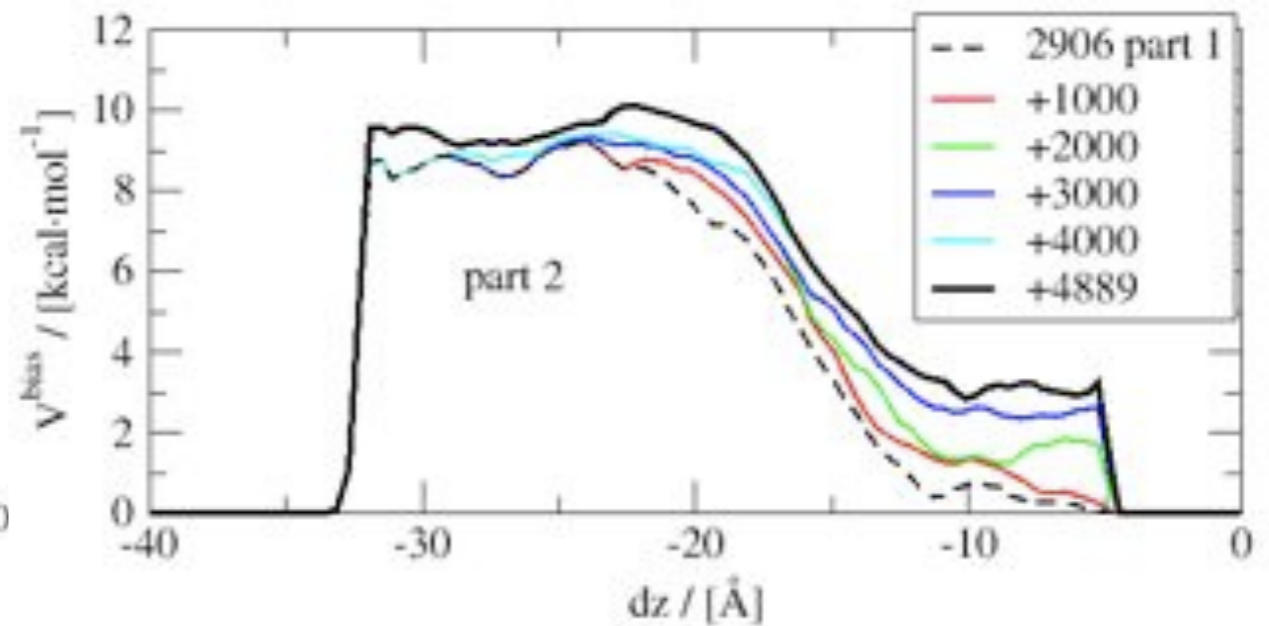
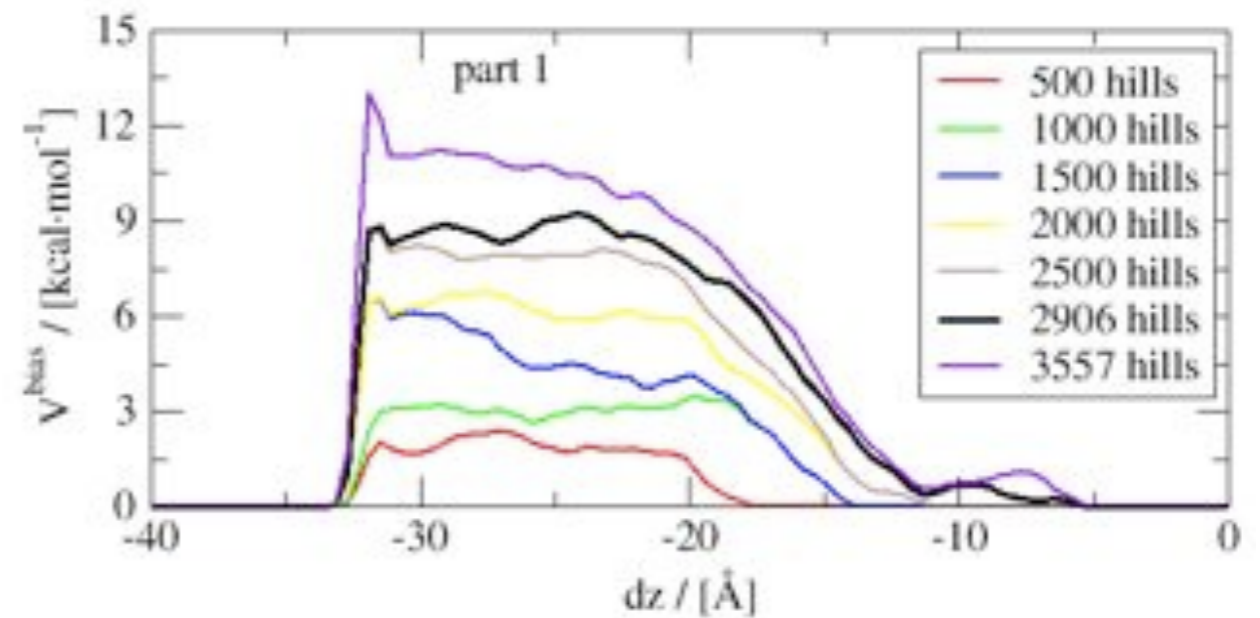
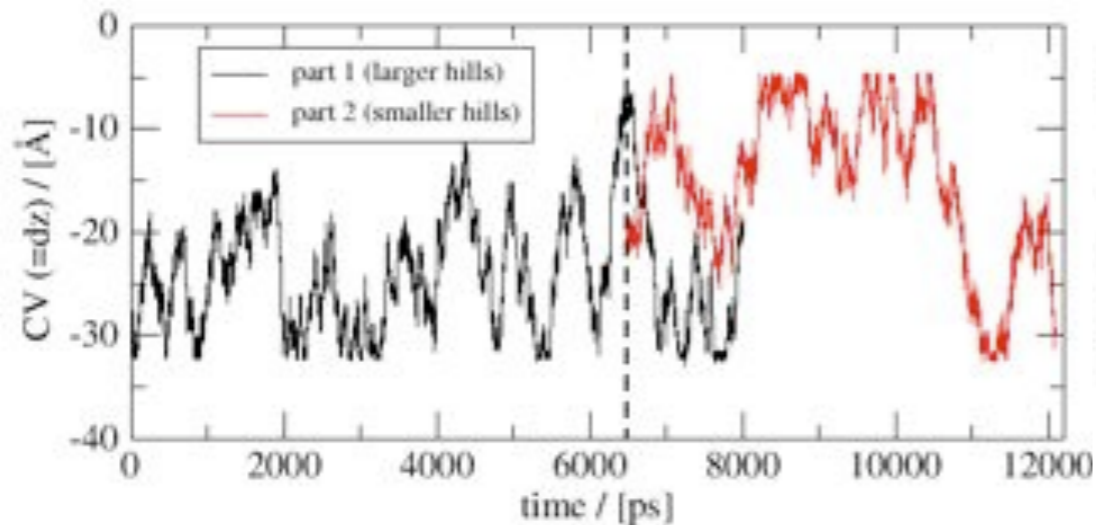
Steered MD



Metadynamics

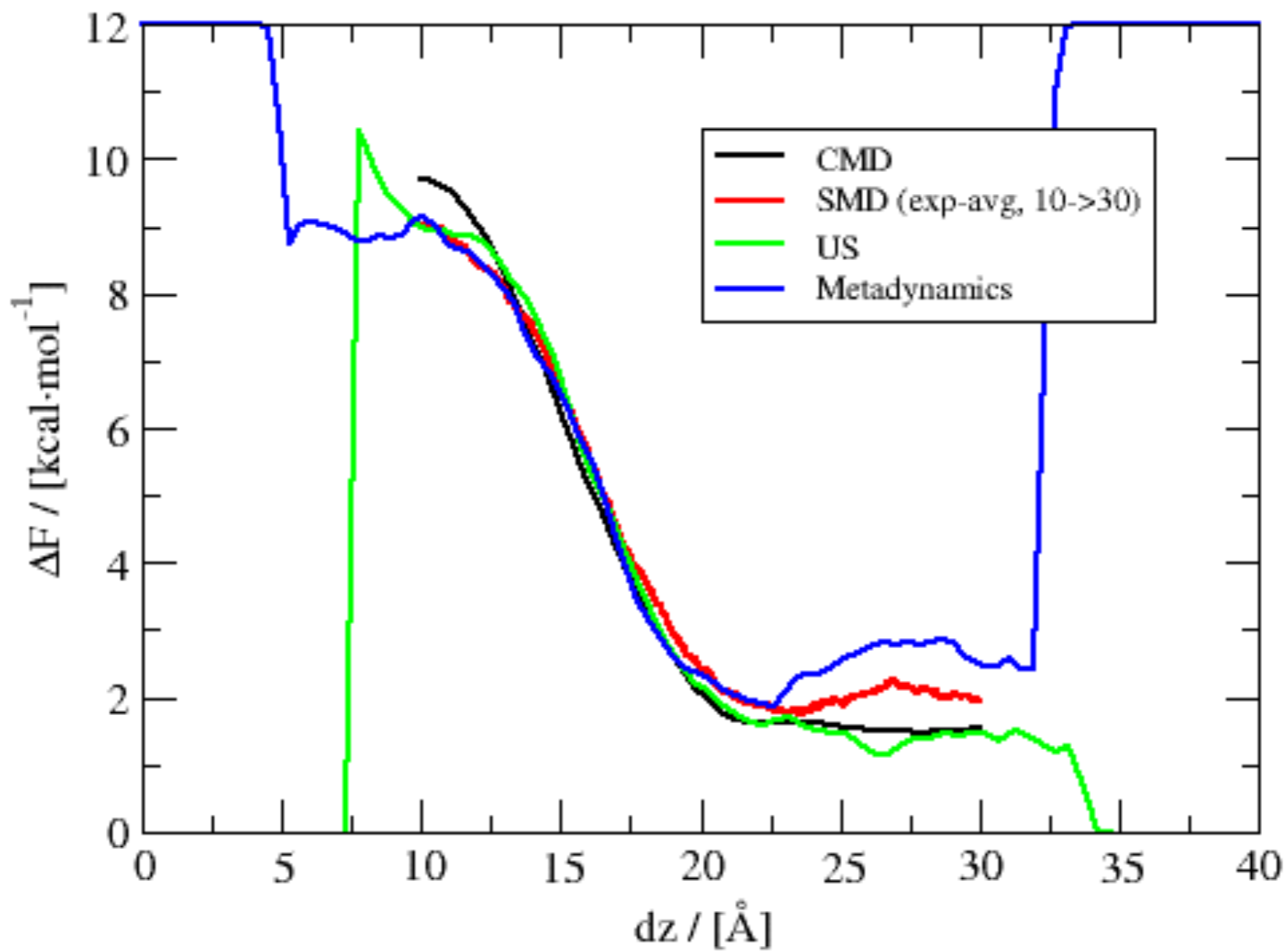
Simulation in two parts with different size of the Gaussian “hills”:

- Part 1: $H=0.25$ K, $W=0.4$ Å
- Part 2: $H=0.10$ K, $W=0.2$ Å



Free energy simulations

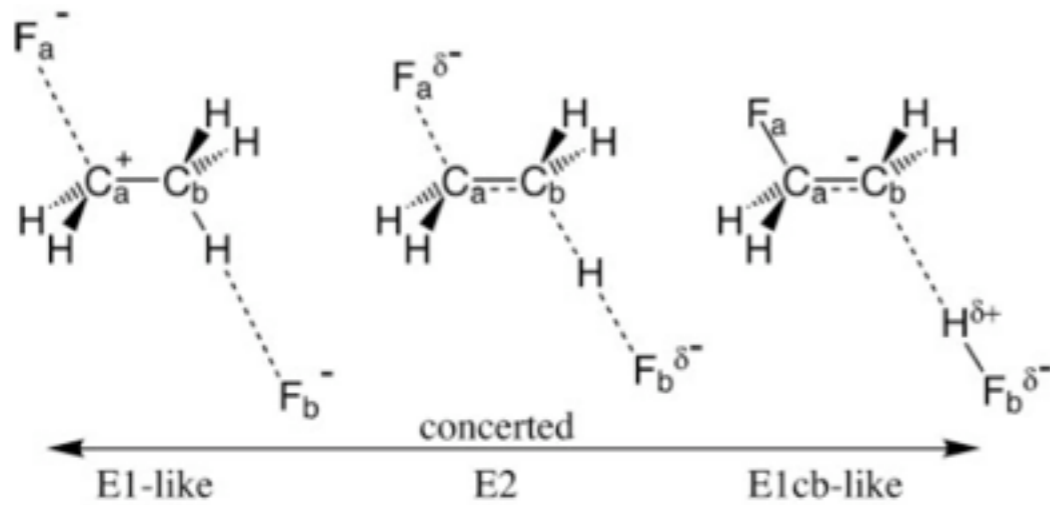
hexane molecule in hexane-water interface



metadynamics

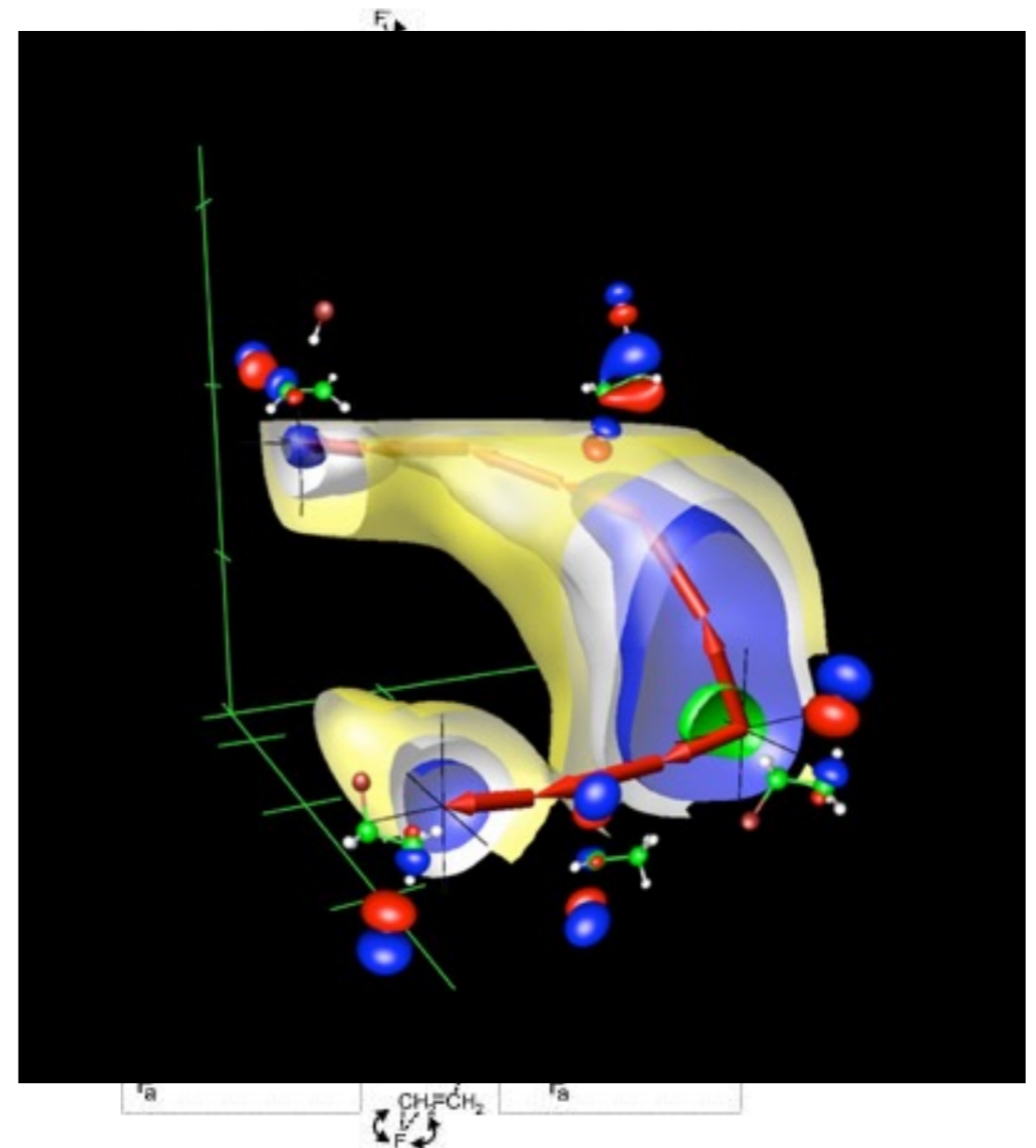
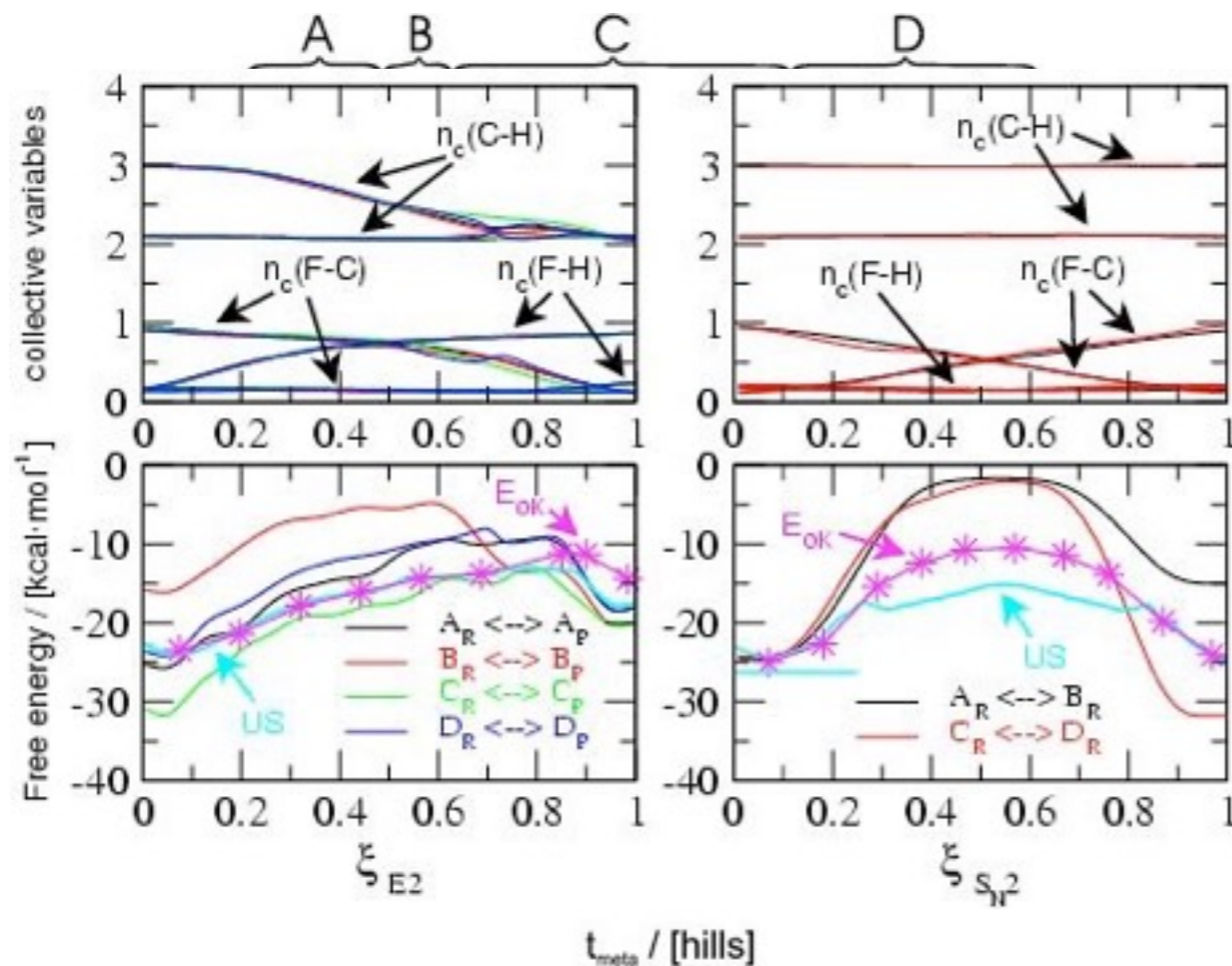
E2 versus Sn2 reaction

(with many collective variables)



8 3D projections

1D collective variables



What about a path collective variable?

Escaping free energy minima

Laio A. and Parrinello M.

Proc. Natl. Acad. Sci. USA **99**, 12562–12566 (2002)

$$V(t, s) = \sum_{t' < t} H_{t'} \prod_{\alpha} \exp \left[\frac{-(s_{\alpha} - s_{\alpha}^{t'})^2}{2\delta_{\alpha}^2 W^2} \right]$$

String method in collective variables:

Minimum free energy paths and isocommittor surfaces

Maragliano L., Fischer A., Vanden-Eijnden E., and Ciccotti G.

J. Chem. Phys. **125**, 024106 (2006)

From A to B in free energy space

Branduardi D., Gervasio F.L., and Parrinello M.

J. Chem. Phys. **126**, 054103 (2007)

$$s(\mathbf{r}) = \frac{1}{P-1} \frac{\sum_{i=1}^P (i-1) e^{-\lambda(\mathbf{r}-\mathbf{r}(i))^2}}{\sum_{i=1}^P e^{-\lambda(\mathbf{r}-\mathbf{r}(i))^2}}$$

$$z(\mathbf{r}) = -\frac{1}{\lambda} \ln \left(\sum_{i=1}^P e^{-\lambda(\mathbf{r}-\mathbf{r}(i))^2} \right)$$

path-metadynamics

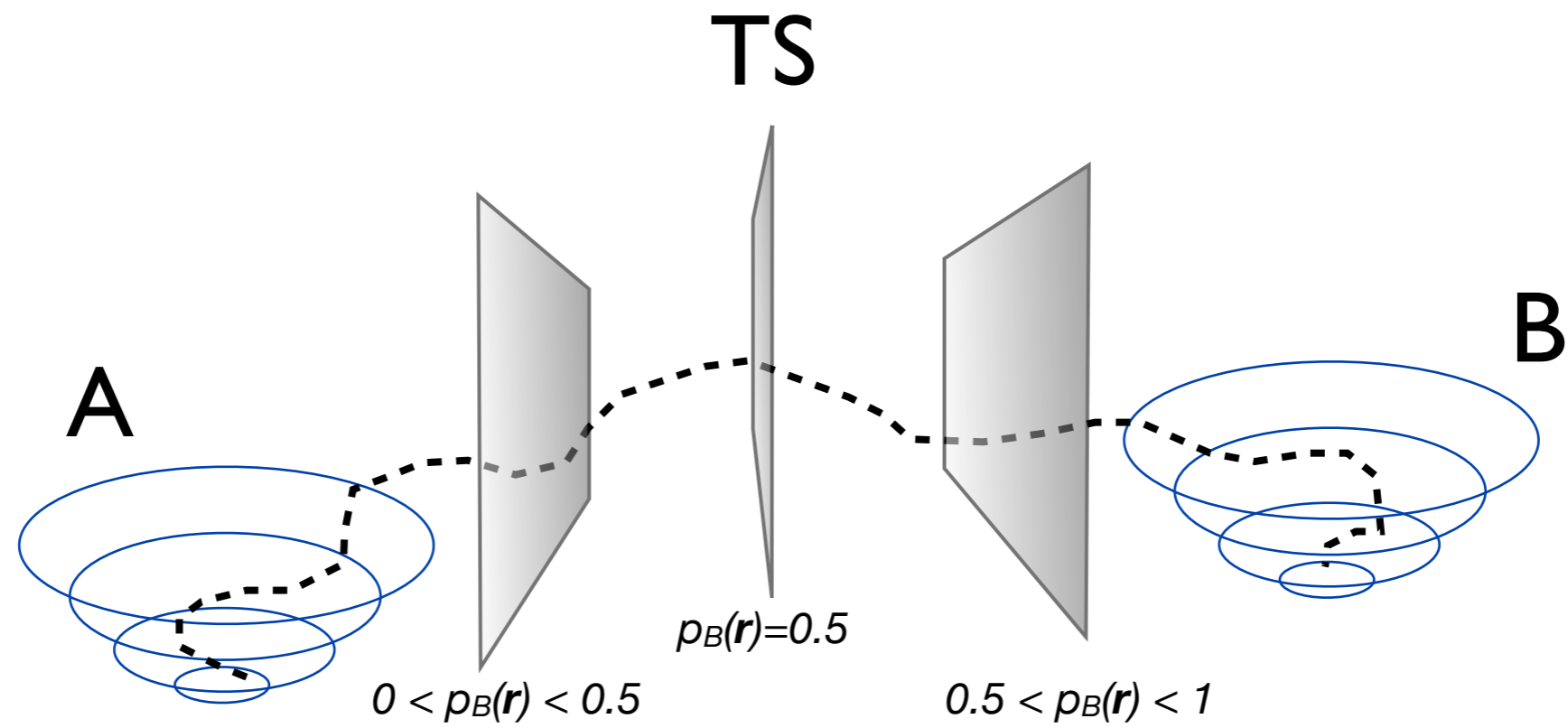
outline

- average transition path or minimum free energy path?
- path: adaptive function of CVs
- simultaneous convergence of path and free energy

- application 1: alanine dipeptide
- application 2: photoactive yellow protein
(see talk of Jocelyne Vreeede)

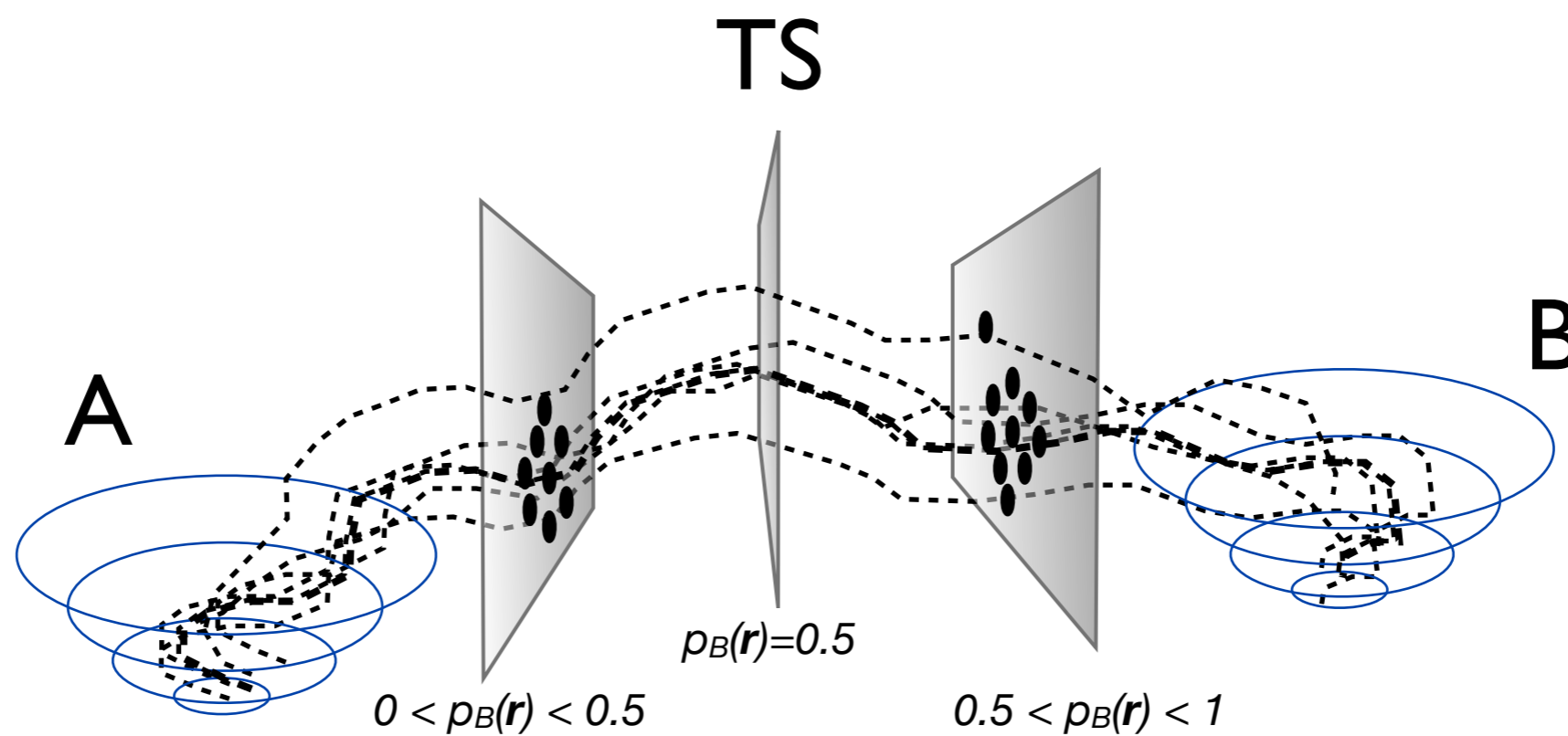
path-metadynamics

iso-committor surfaces



path-metadynamics

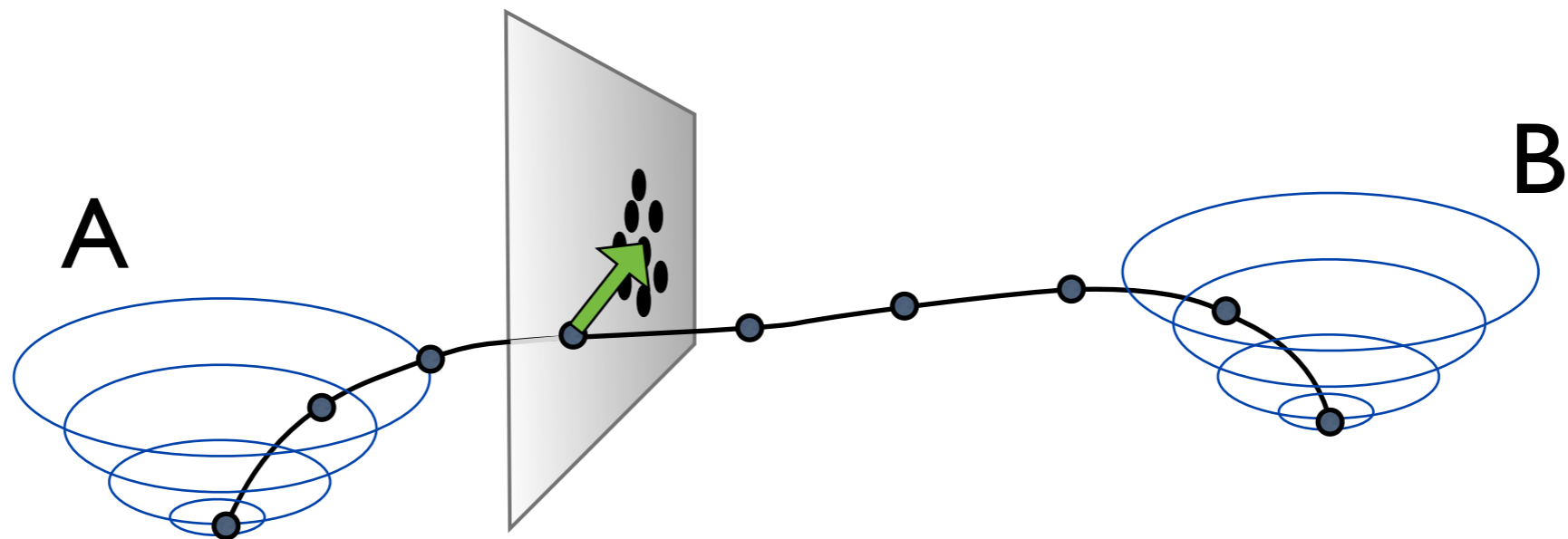
iso-committor surfaces



the average transition pathway
(in CV space)

path-metadynamics

distance to mean density



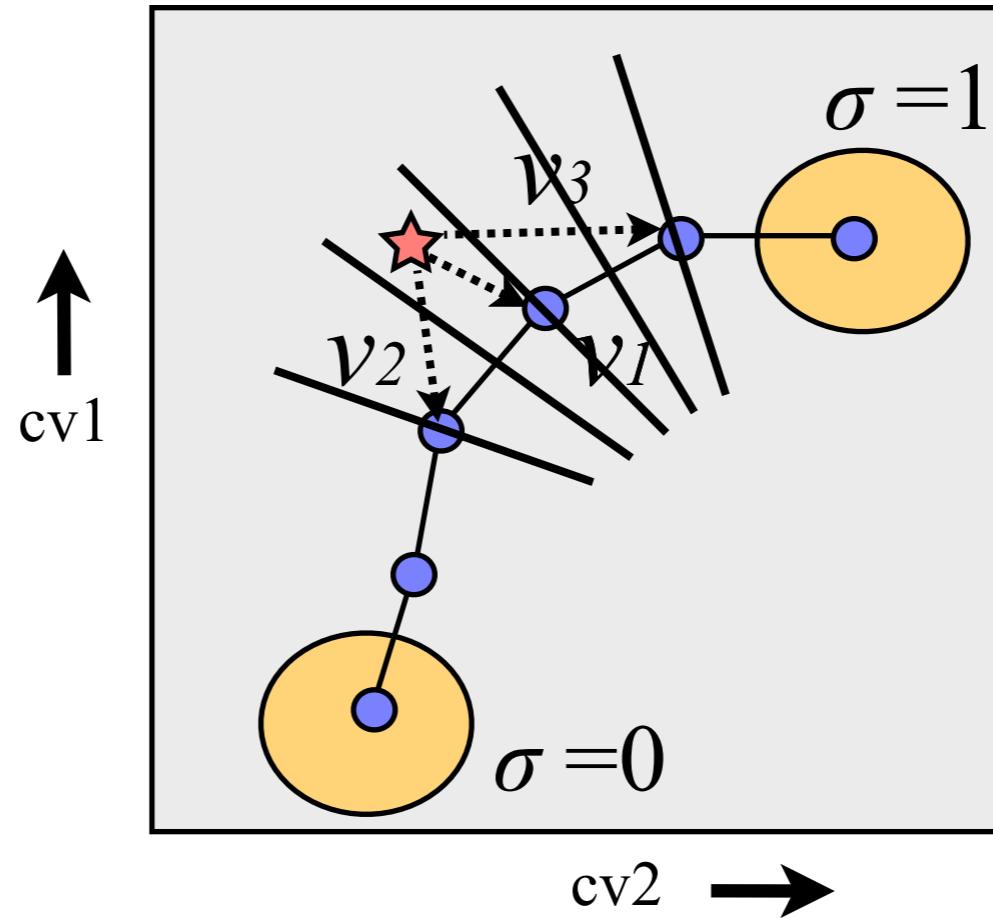
- start from guess path
- bias dynamics along path
- move nodes to the mean density
- maintain equidistant nodes

practical implementation

adaptive collective variable

- add extra collective variable: σ
- σ is a function of all other coll. variables
- biasing potential is only working on σ
- σ function adapts on the fly

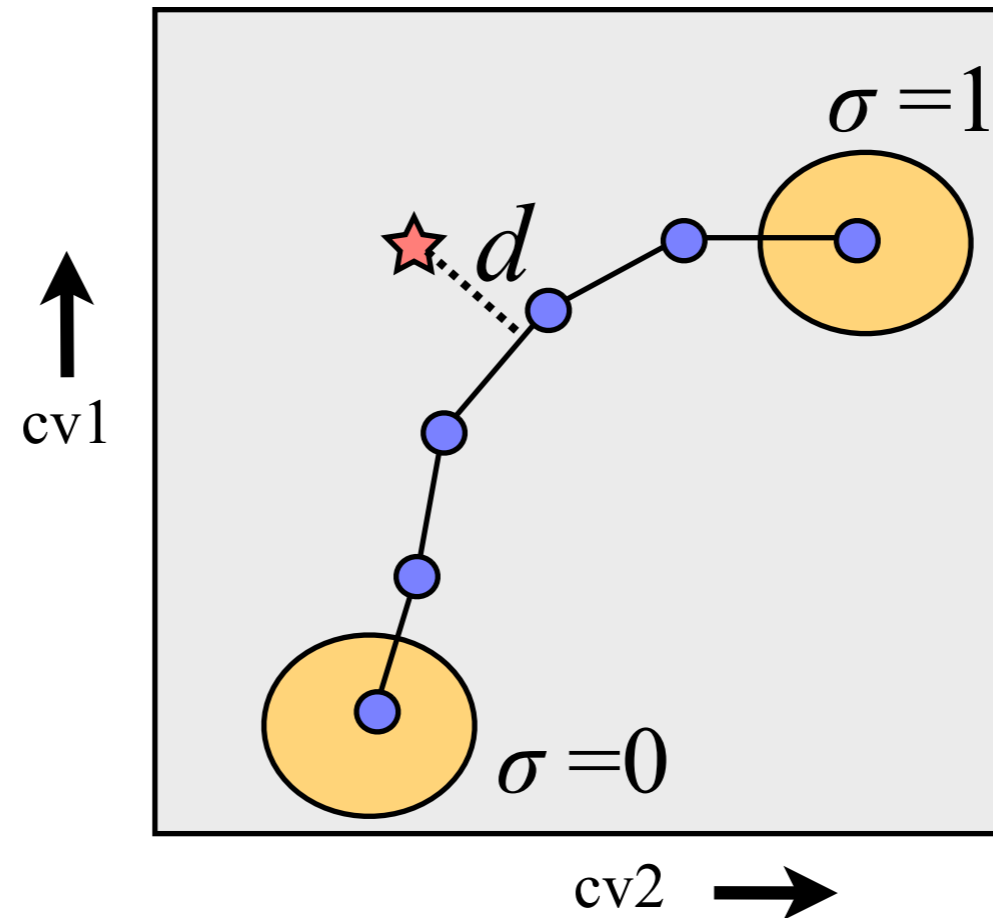
practical implementation



path collective variable

$$\sigma \equiv \frac{i_{min}}{n} \left(i_{min} \pm \frac{\sqrt{(\bar{v}_1 \bar{v}_3)^2 - \bar{v}_3 \bar{v}_3 (\bar{v}_1 \bar{v}_1 - \bar{v}_2 \bar{v}_2) - \bar{v}_1 \bar{v}_3}}{2(\bar{v}_3 \bar{v}_3 - 1)} \right)$$

practical implementation



distance from path

histogram distance to the path for each node
maximum of $P(d)$ should be at $d=0$

or accumulate average d for each node

practical implementation

leverage measurement of d
between closest nodes

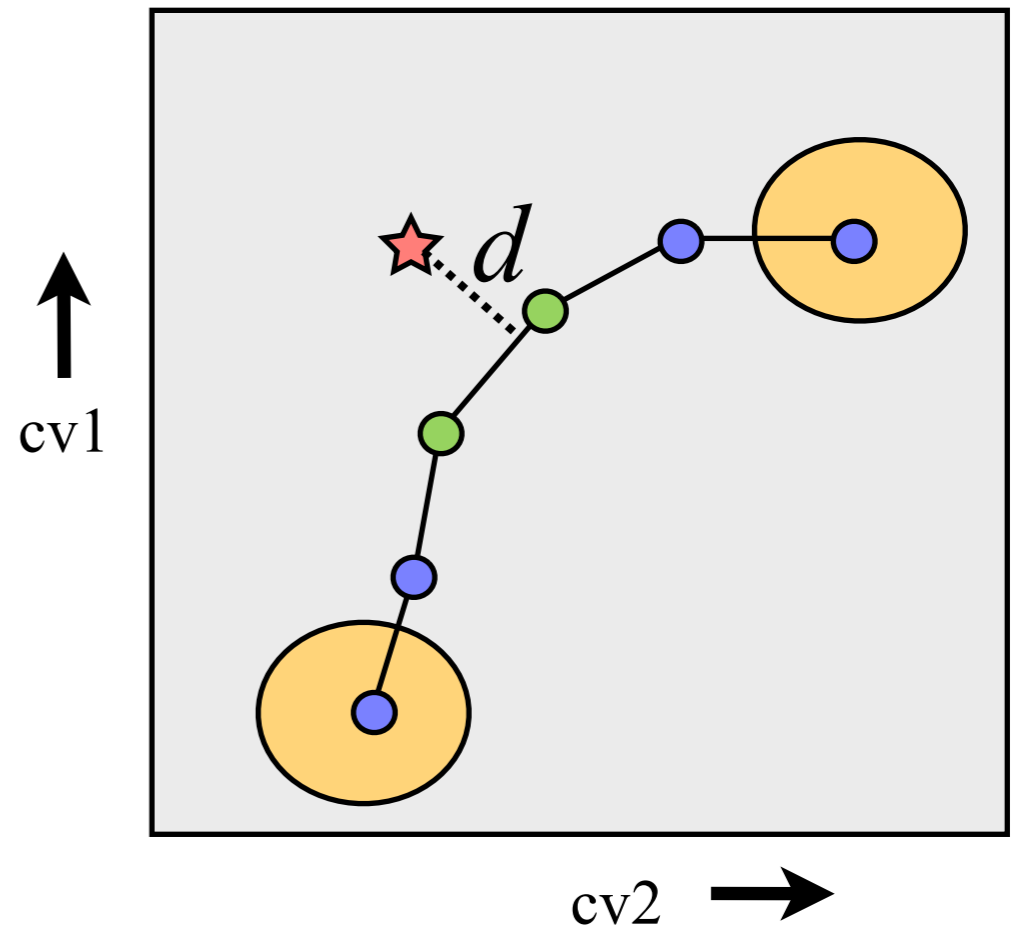
$$\mathbf{s}_j^{t_{i+1}} = \mathbf{s}_j^{t_i} + \sum_k w_k \cdot |\mathbf{s}^{t_i}(\sigma(\mathbf{z}_k)) - \mathbf{z}_k| / \sum_k w_k$$
$$w_k = \max\left(0, \left(1 - \frac{|\mathbf{s}_j^{t_i} - \mathbf{s}^{t_i}(\sigma(\mathbf{z}_k))|}{|\mathbf{s}_j^{t_i} - \mathbf{s}_{j+1}^{t_i}|}\right)\right)$$

path update

move nodes to d (every step)
set $d=0$
redistribute nodes along path

half life time of data

$$f = \exp[\tau^{-1} * \ln \frac{1}{2}]$$

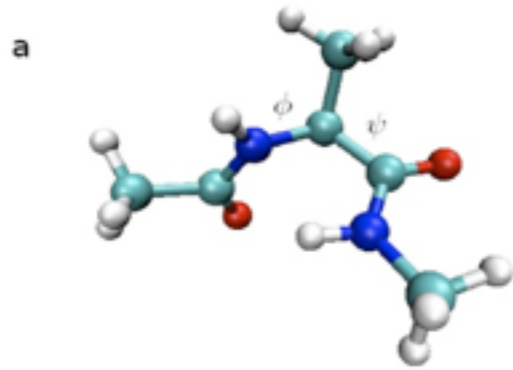


optional tube potential

$$V^{\text{tube}} = \frac{k}{2} d^2 .$$

alanine dipeptide

• Zhiwei Liu
• Preston Moore
University of the Sciences, Philadelphia



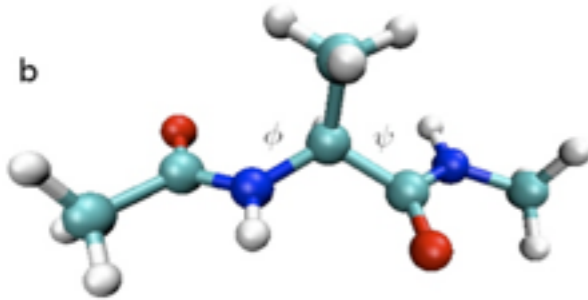
Classical Molecular Dynamics
(CM3D code by Preston Moore, USP)

Alanine dipeptide + 216 water
CHARMM27 forcefield (modified internal params water)
Cubic periodic box L=18.8 Angstrom
NVT ensemble, T=298K

Hill size: W=0.2 rad, H=0.02 kcal/mol
Hill stride: 100 MD steps (=50 fs)

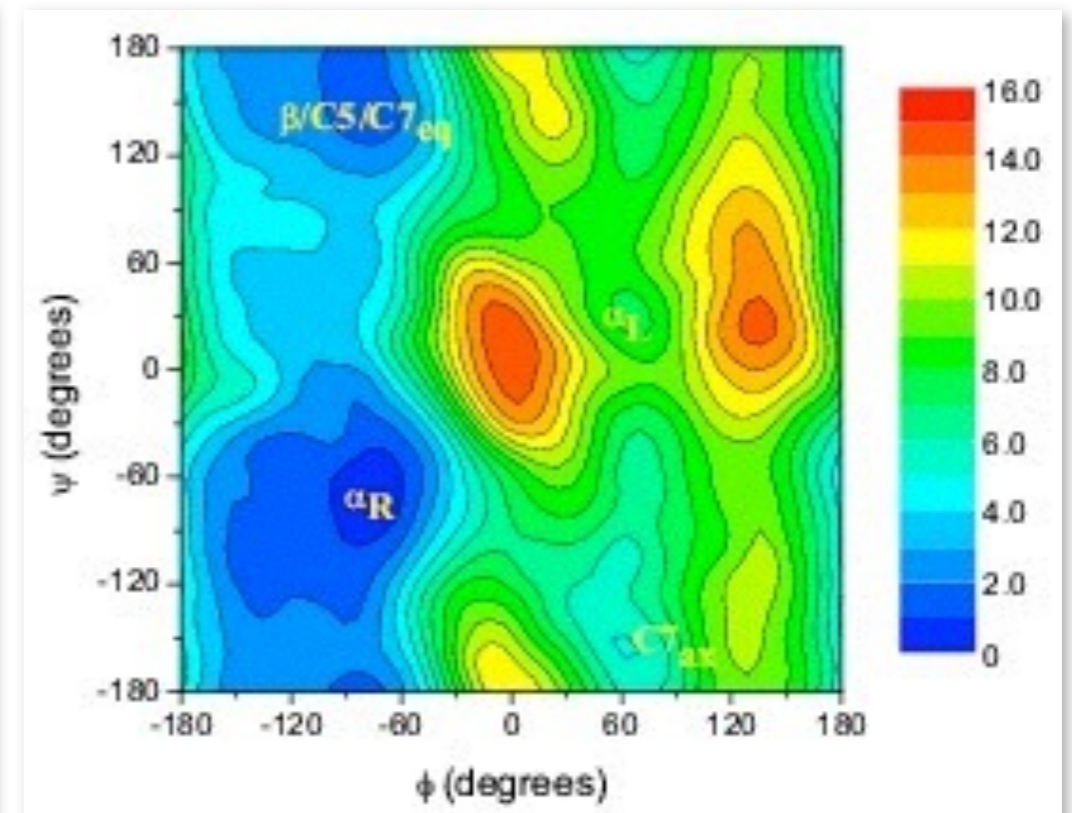
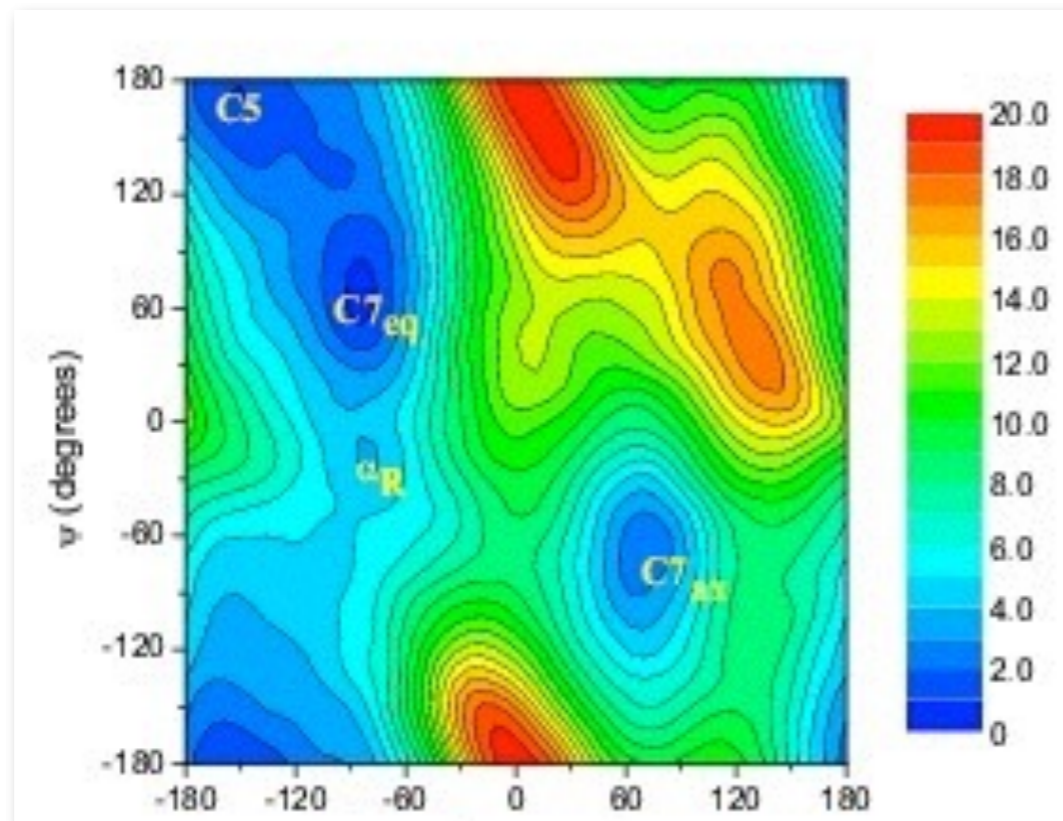
Total simulation time: ca. 5 ns

Error in energy: 0.3-0.5 kcal/mol
Error in angles: 3-10 degrees



Gas-phase

Aqueous solution



alanine dipeptide

Parameters:

$$T = 300 \text{ K}$$

$$H_{\text{gaussian}} = 10 \text{ K}$$

$$W_{\text{gaussian}} = 0.05$$

$$\Delta t_{\text{gaussian}} = 100 \text{ MD steps}$$

$$n = 20 + 20 + 20$$

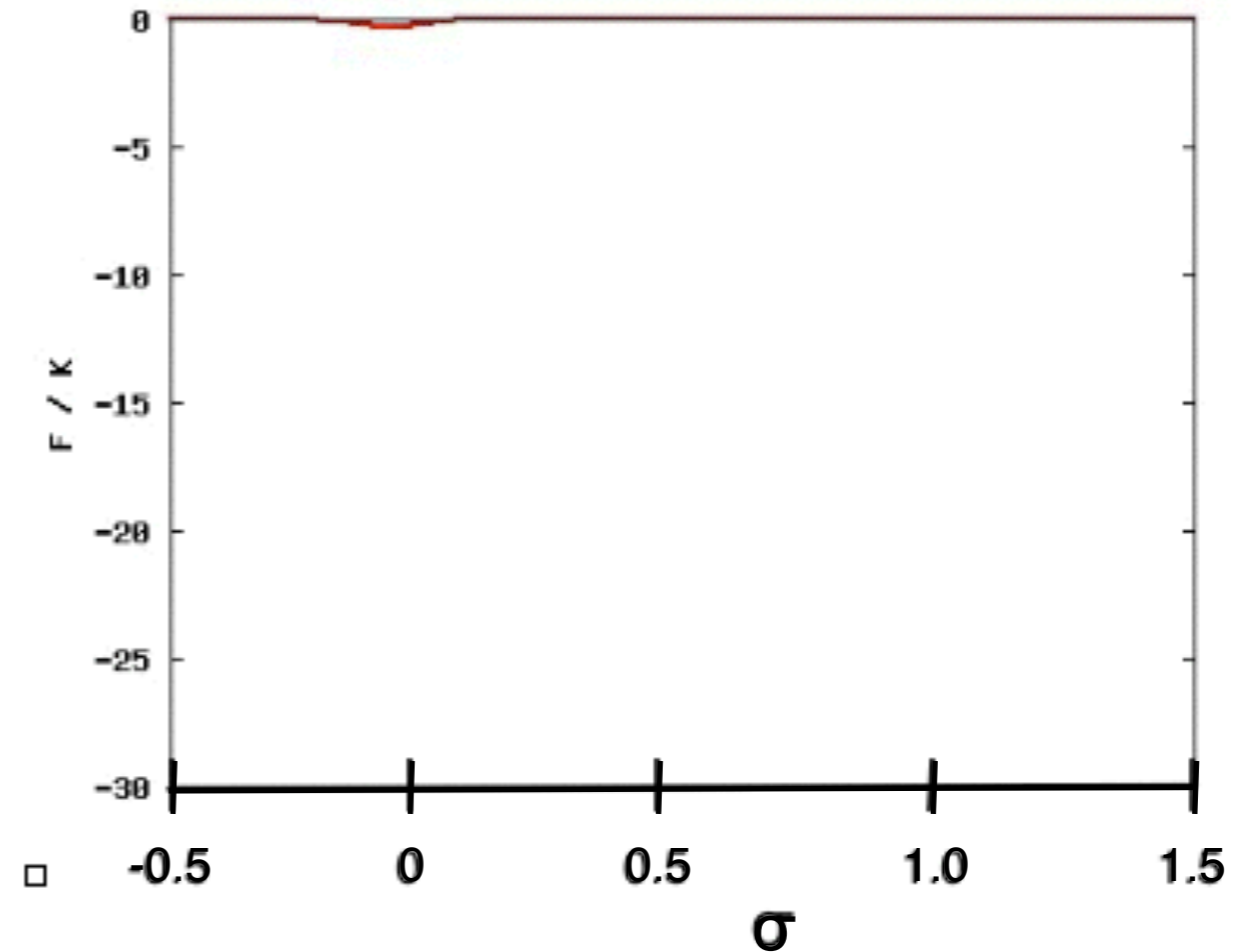
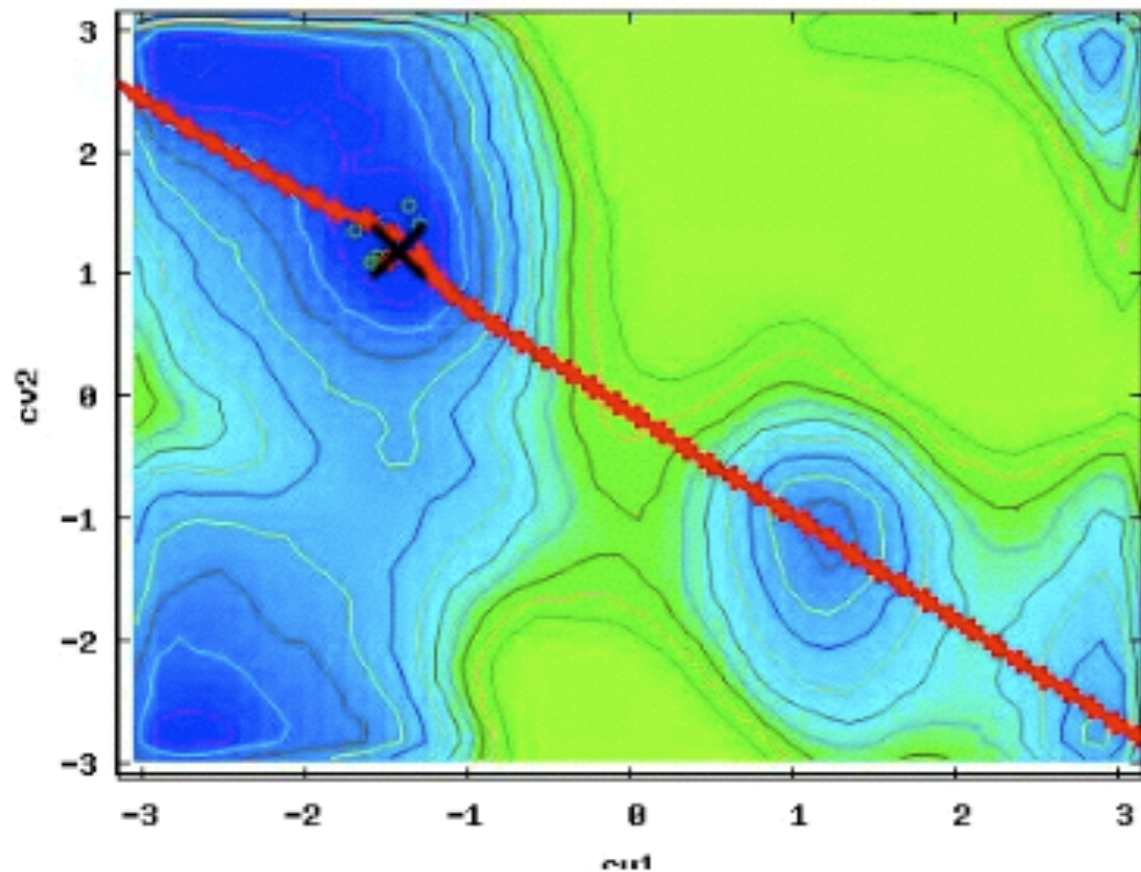
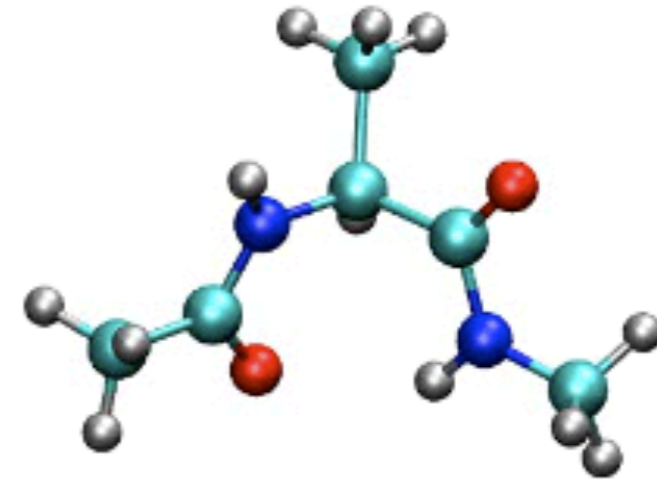
$$\tau = 1000 \text{ MD steps}$$

every recrossing:

$$H \times 50\%$$

$$W \times 50\%$$

$$\tau \times 100$$



alanine dipeptide

Parameters:

$$T = 300 \text{ K}$$

$$H_{\text{gaussian}} = 10 \text{ K}$$

$$W_{\text{gaussian}} = 0.05$$

$$\Delta t_{\text{gaussian}} = 100 \text{ MD steps}$$

$$n = 20 + 20 + 20$$

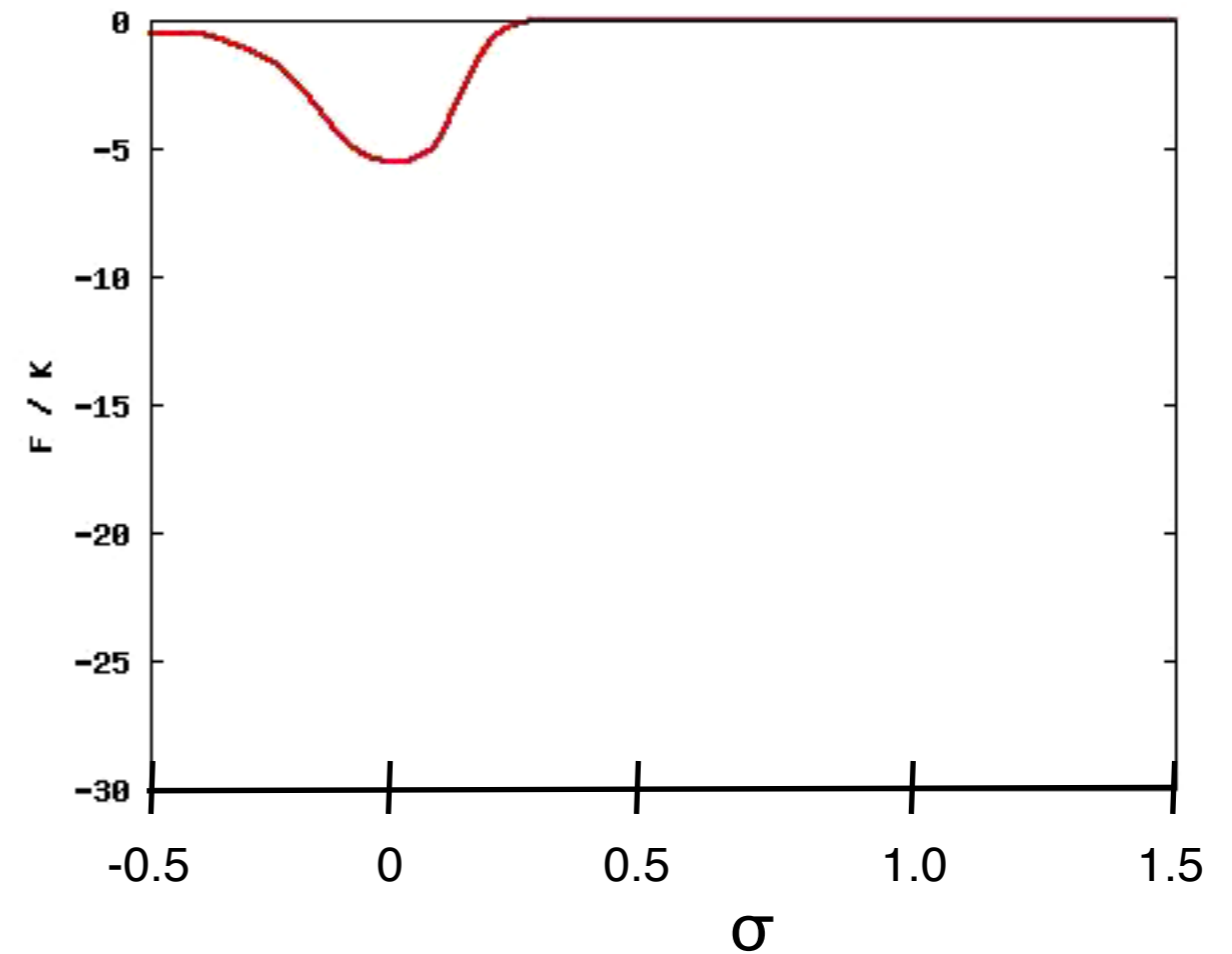
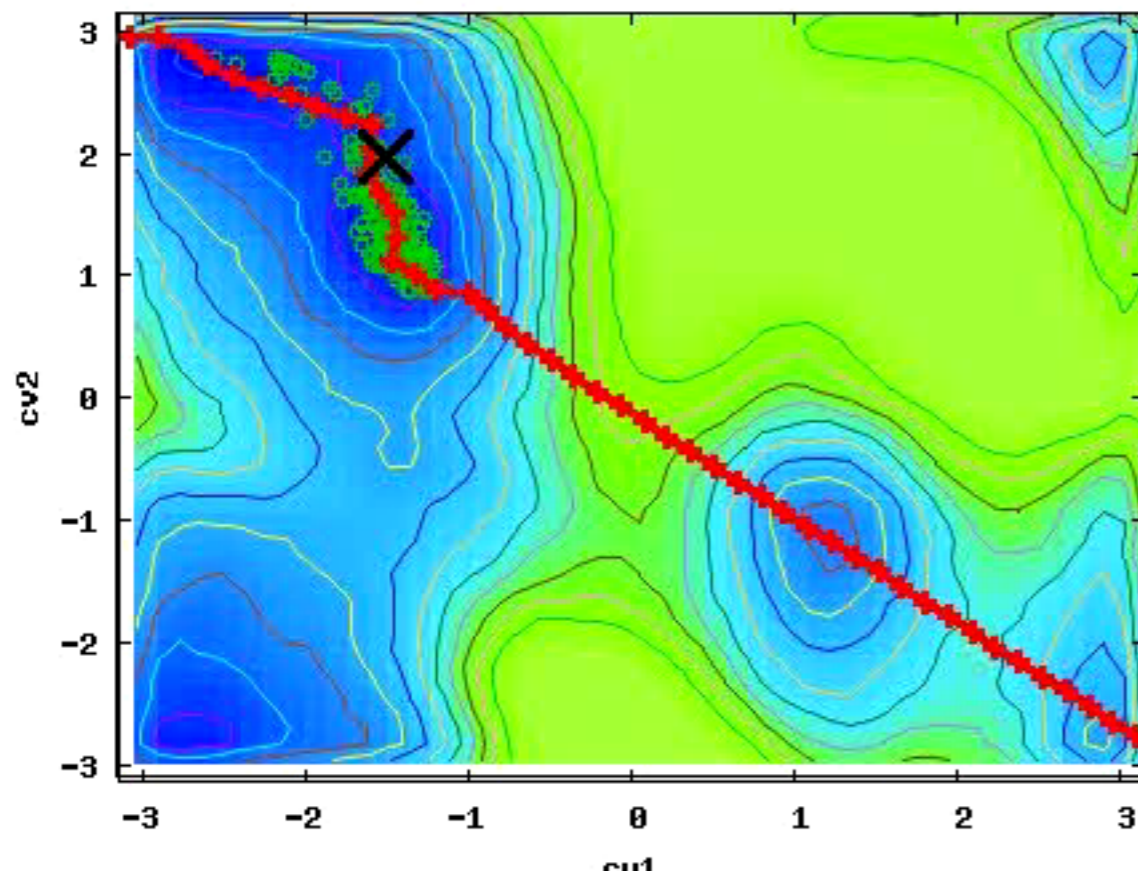
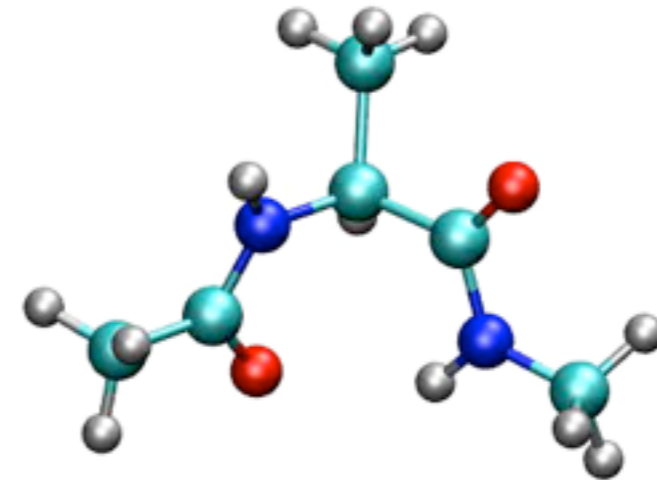
$$\tau = 1000 \text{ MD steps}$$

every recrossing:

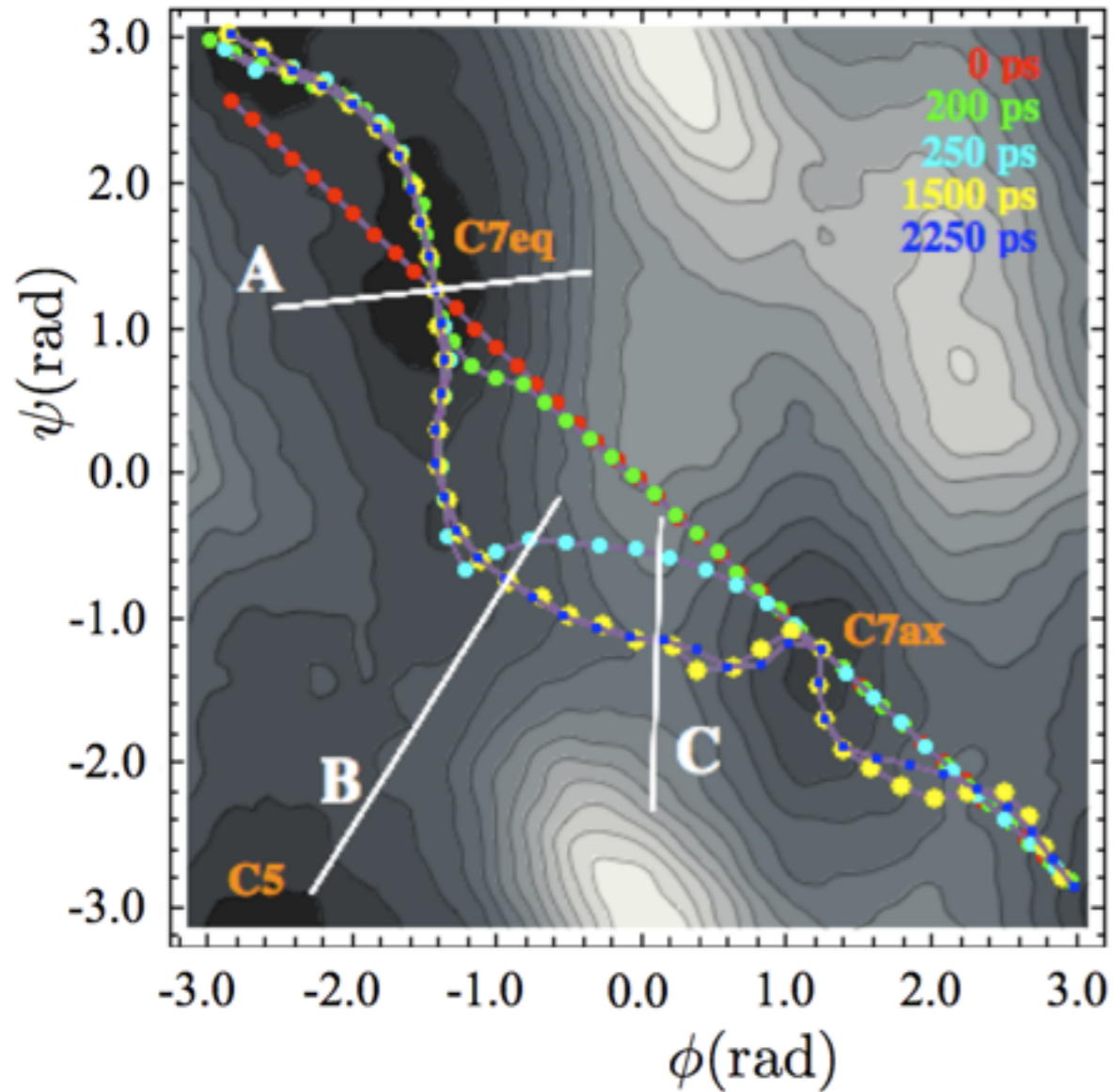
$$H \times 50\%$$

$$W \times 50\%$$

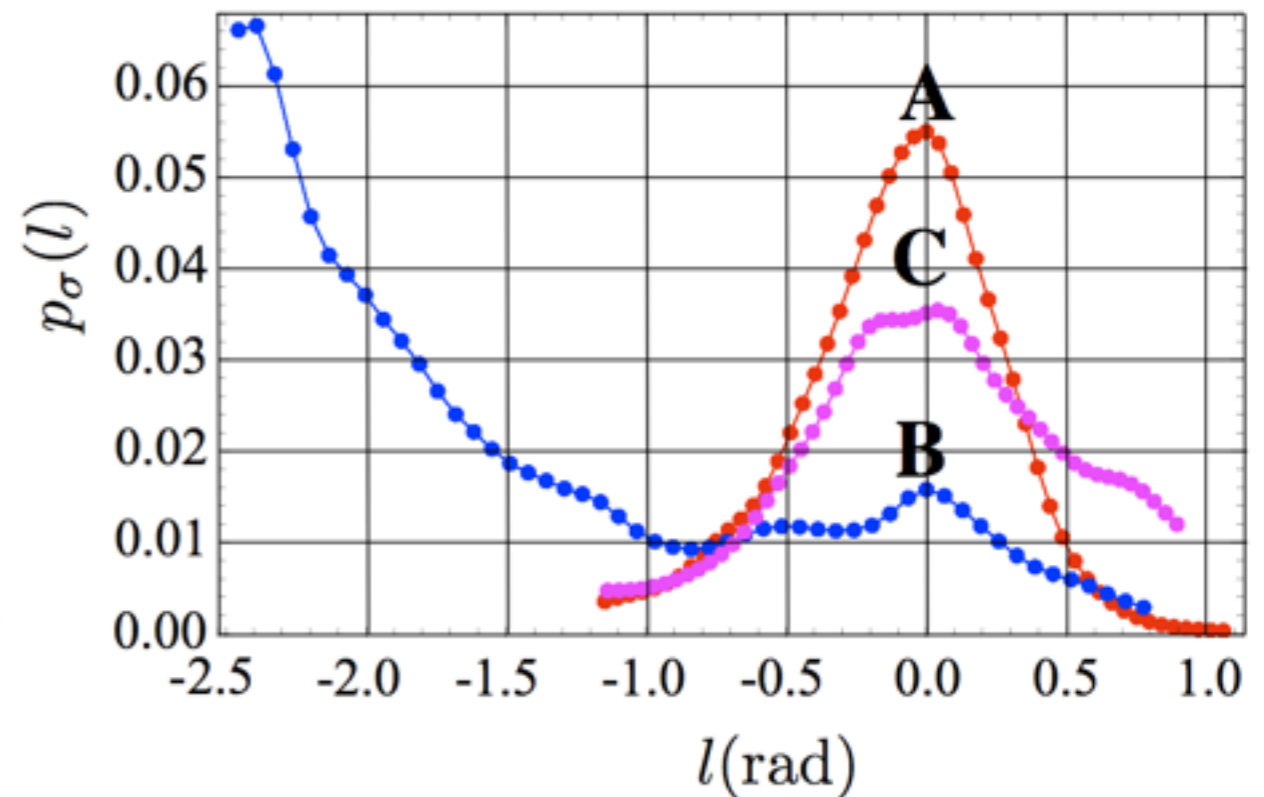
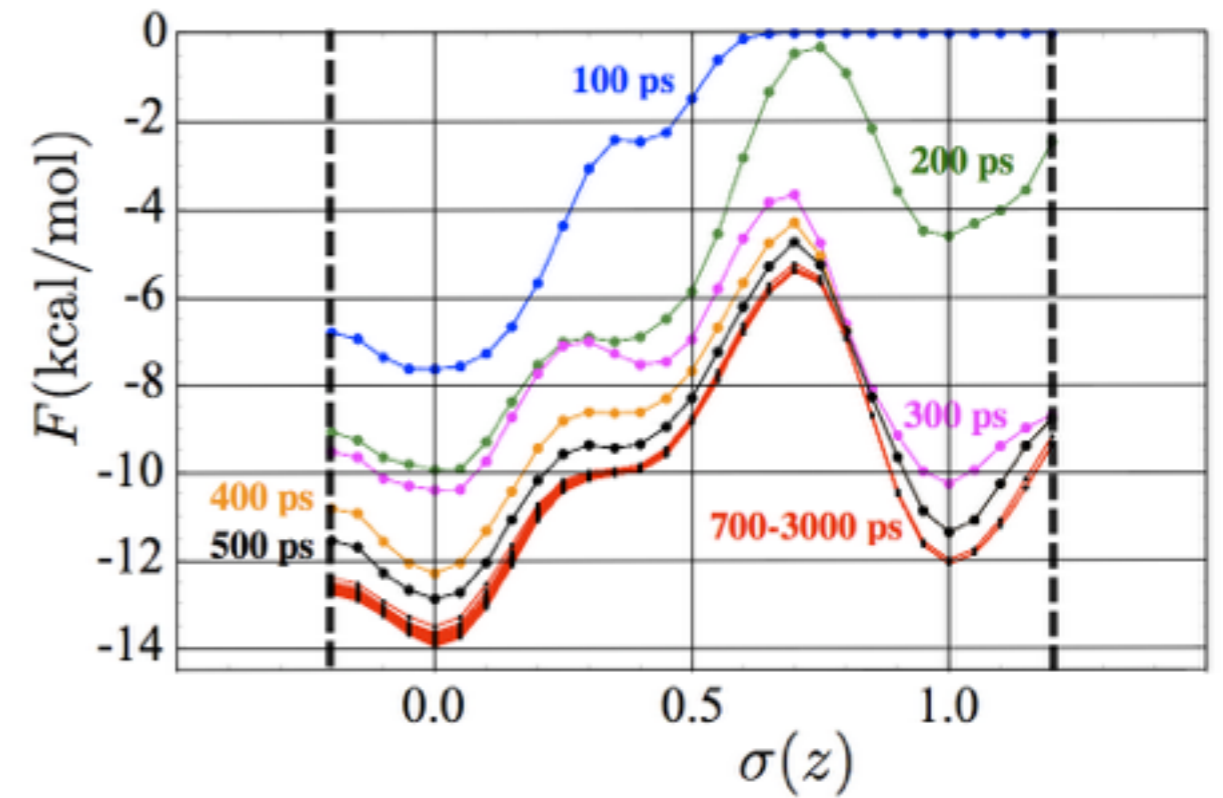
$$\tau \times 100$$



convergence

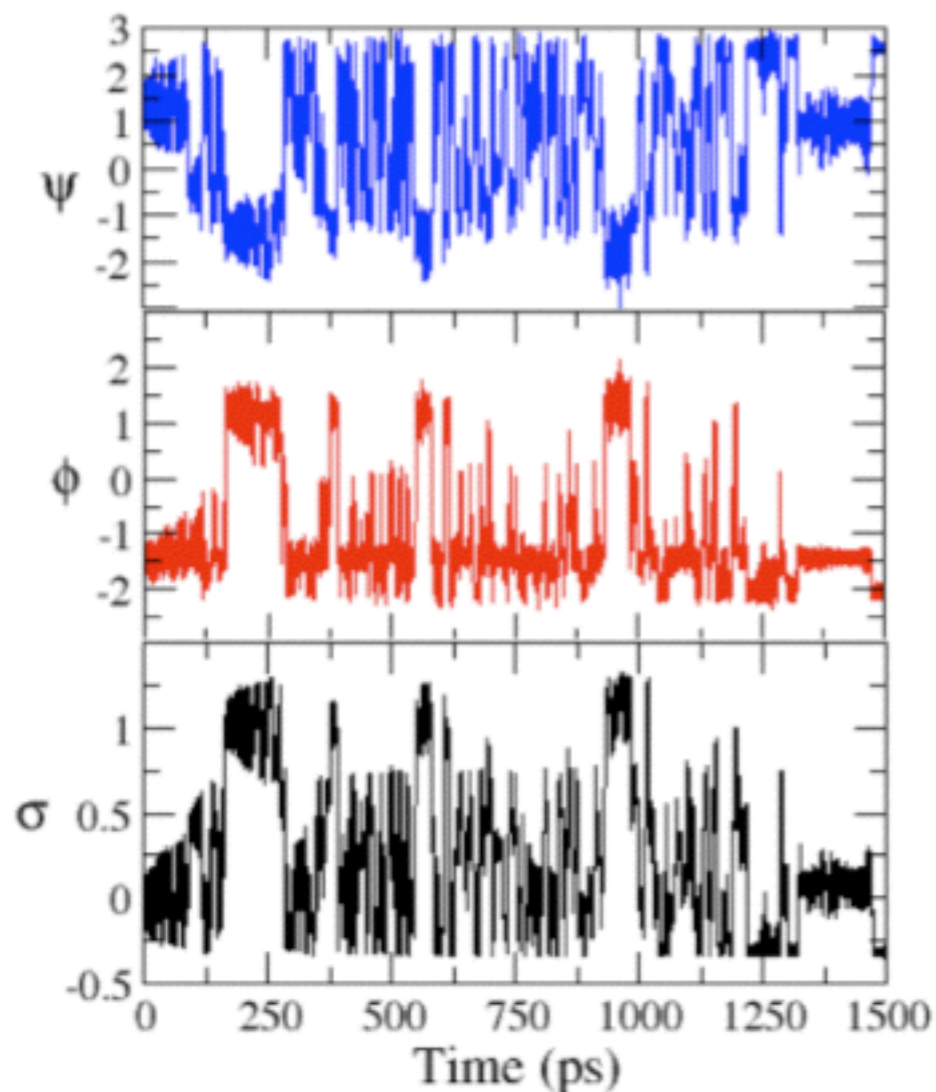


- tube potential was needed
- well-tempered metadynamics also works



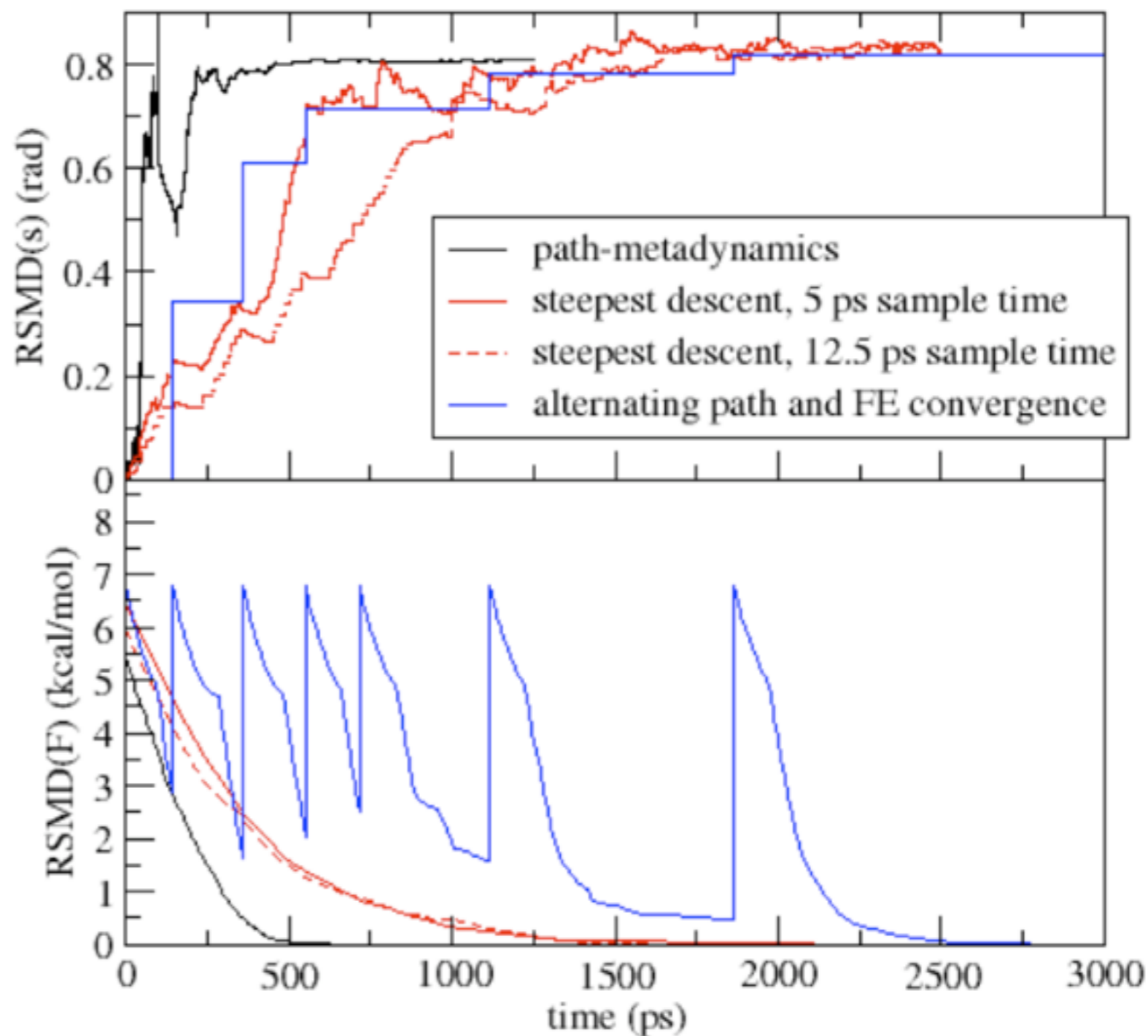
Path Finding on High-Dimensional Free Energy Landscapes
Grisell Díaz Leines and Bernd Ensing,
Phys. Rev. Lett. 109, 020601 (2012)

convergence

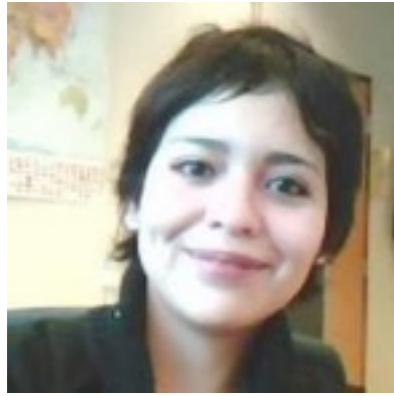


$$\text{RMSD}(s) = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(|\mathbf{s}_i(t) - \mathbf{s}_i(0)| \right)^2}$$

$$\text{RMSD}(F) = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(F_i(t) - F_i^{\text{final}} \right)^2}$$



path-metadynamics



Grisell Diaz Leines



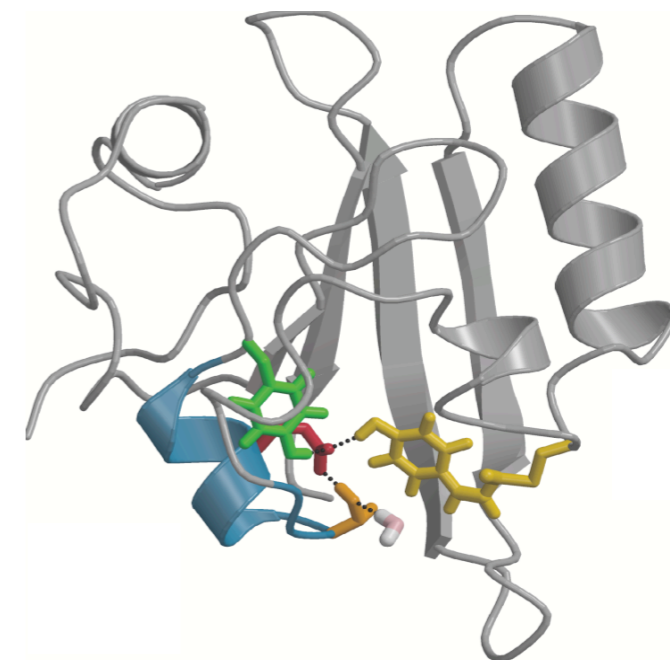
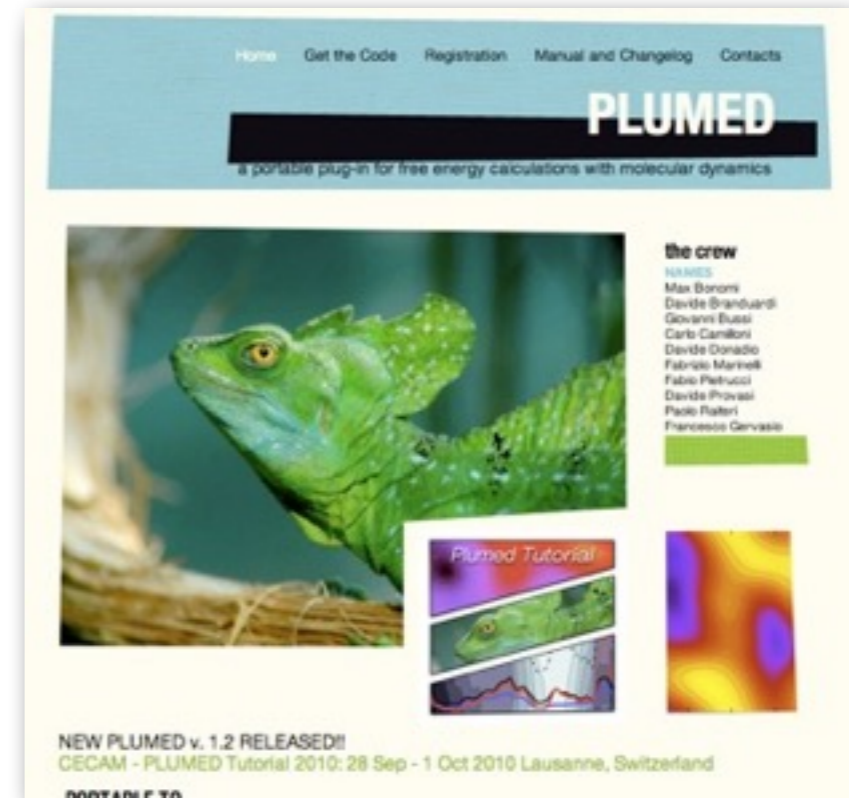
Davide Branduardi



Wagner Brandeburgo

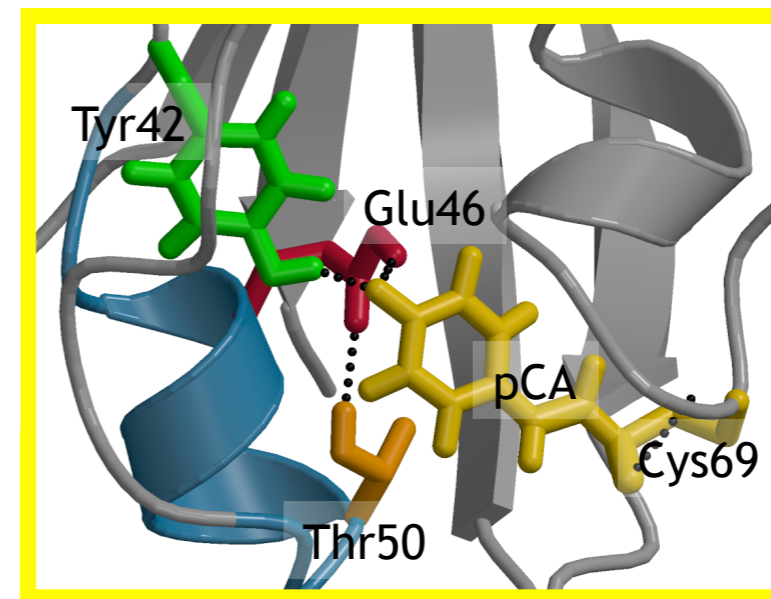
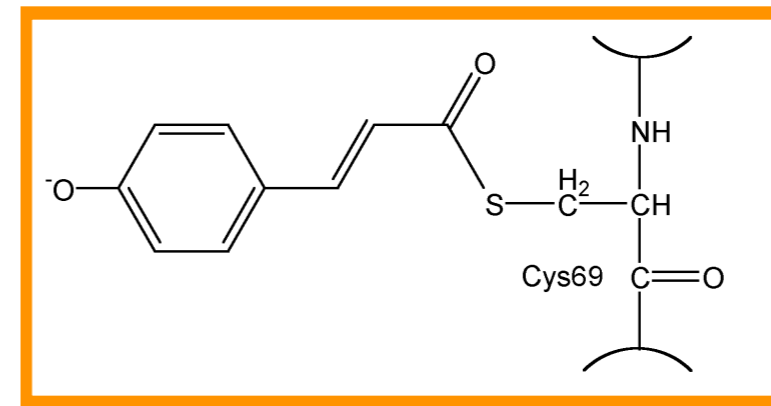
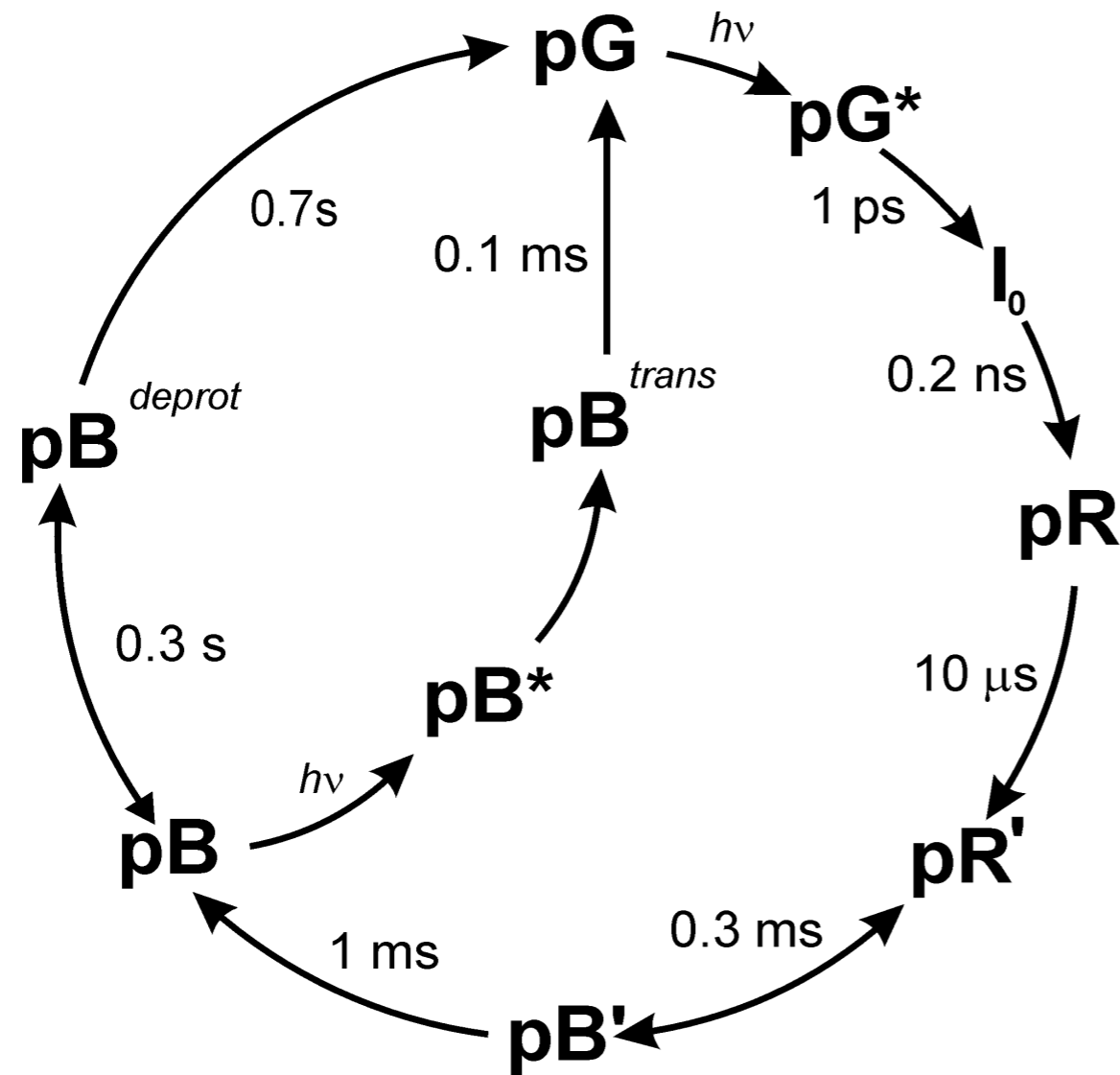


Jocelyne Vreede



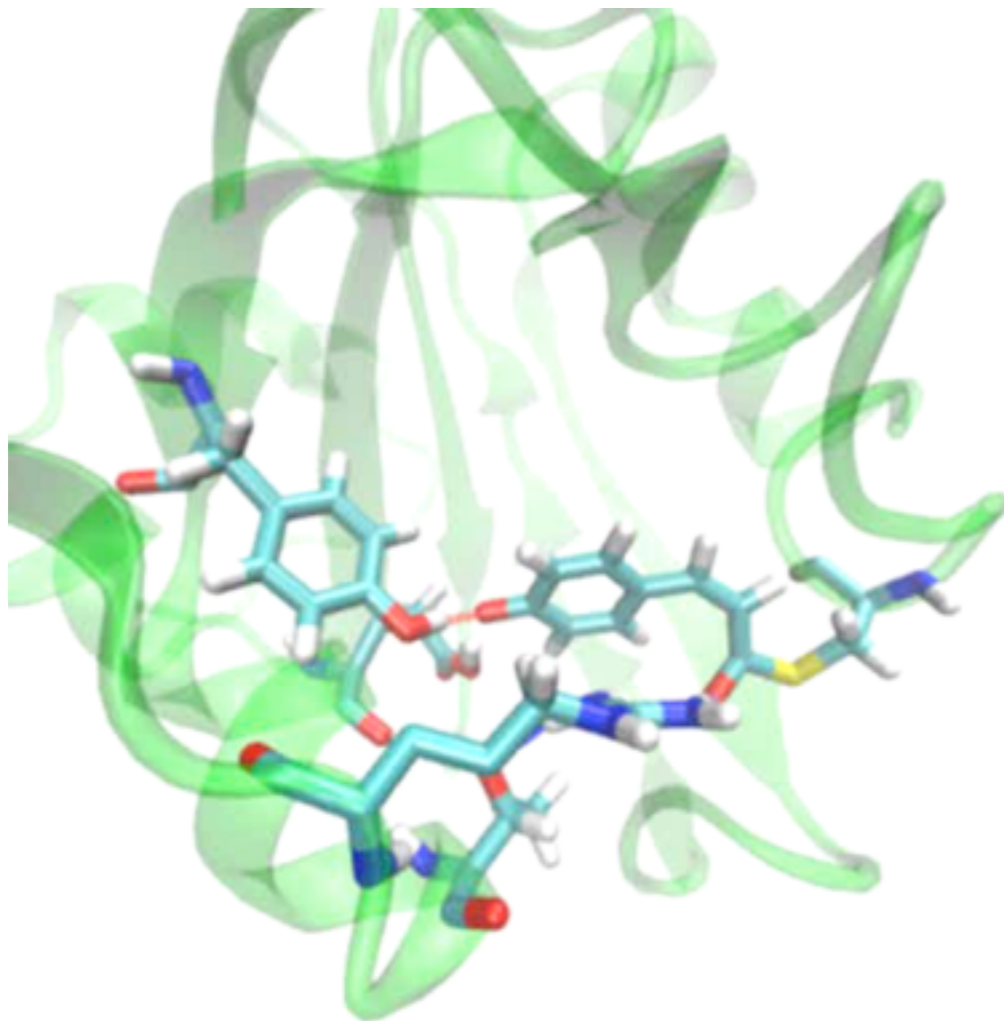
Photoactive Yellow Protein

blue-light sensor from *H. halophila*

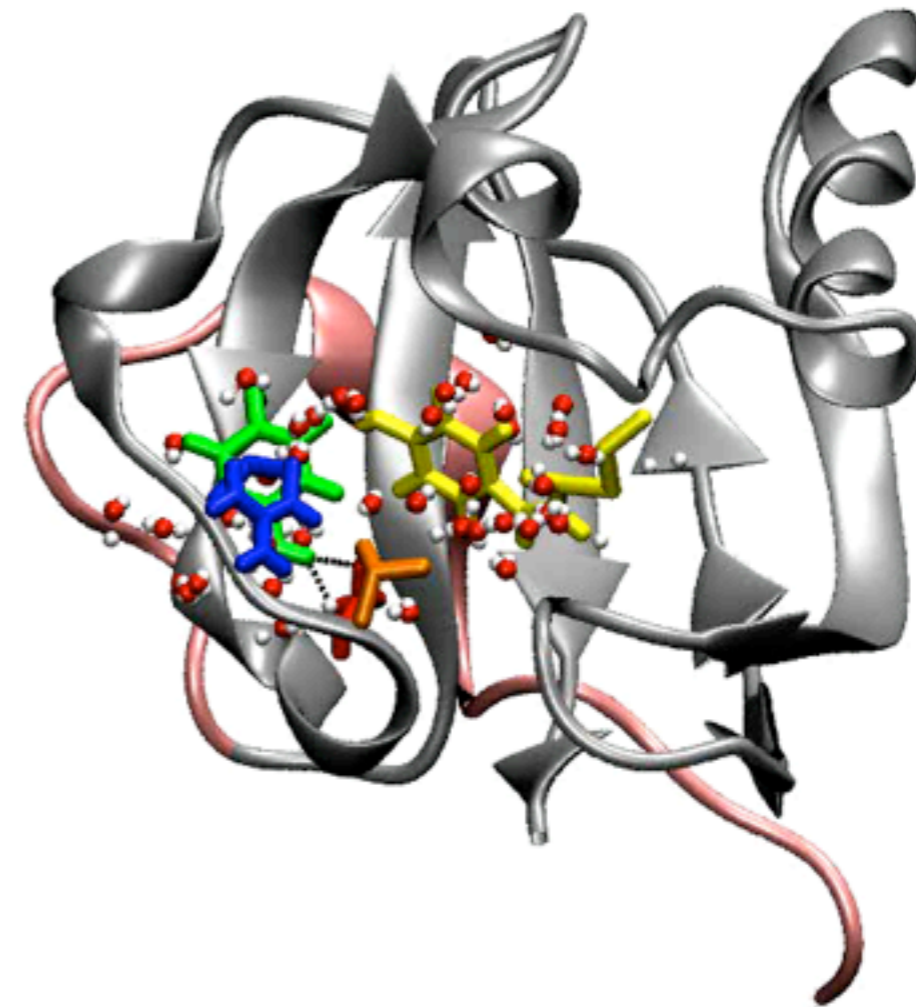


Transition path sampling

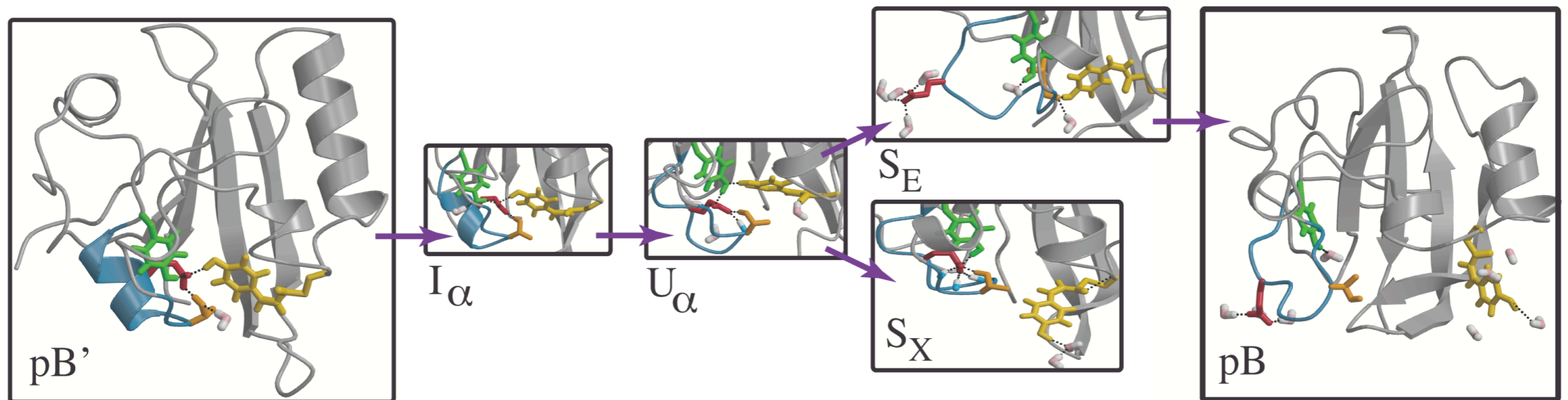
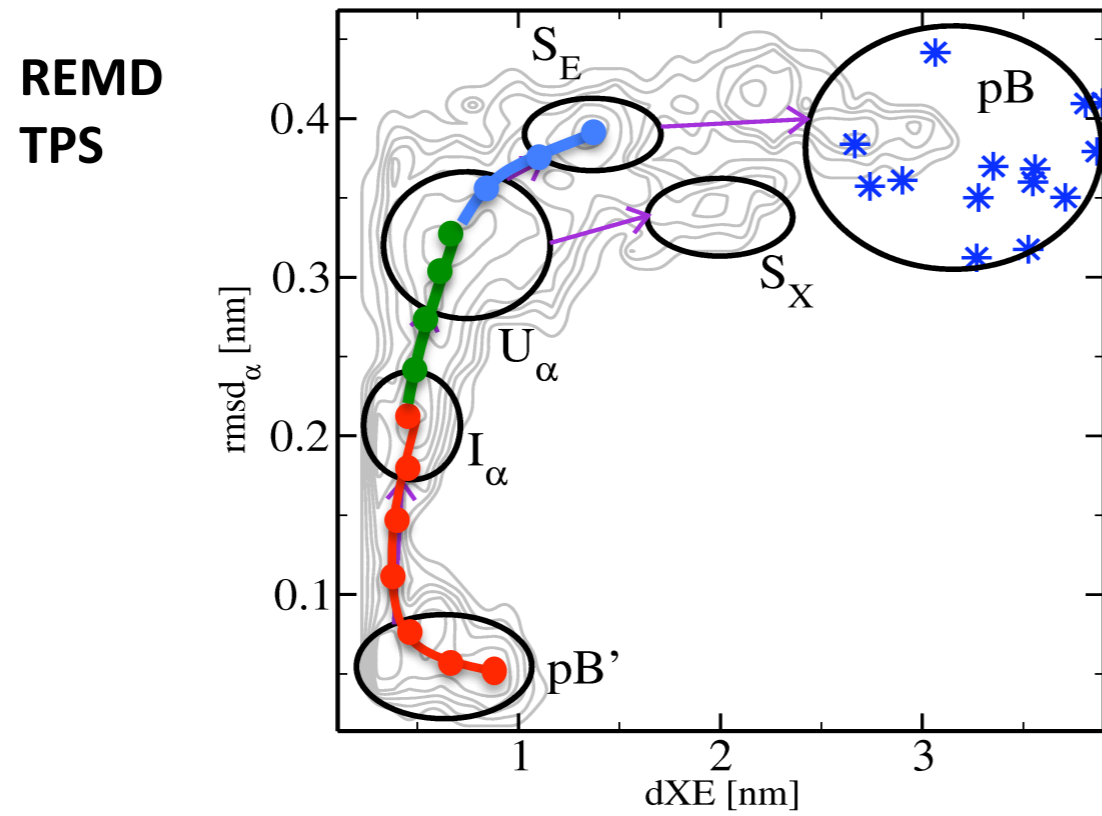
Proton transfer step
Siri van Keulen, Peter Bolhuis, BE



Conformational transformation
Jocelyne Vreede, Peter Bolhuis, et al.



Signaling state



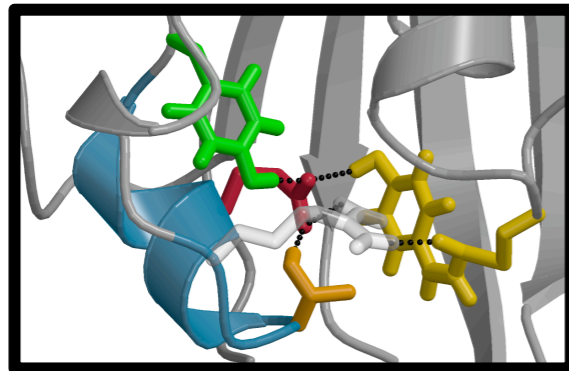
path-metadynamics study of PYP

TPS:

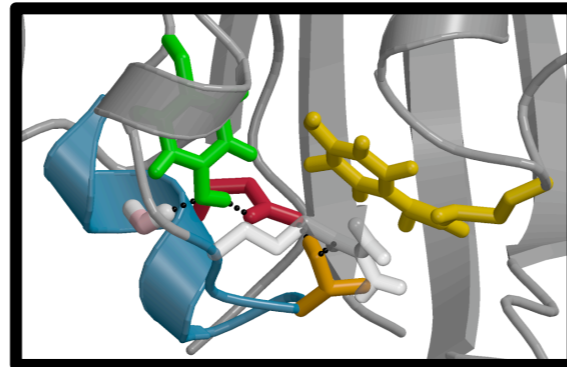
3847 accepted pathways
180 uncorrelated pathways
~1 μ s total simulation time

unfolding of helix 43-51

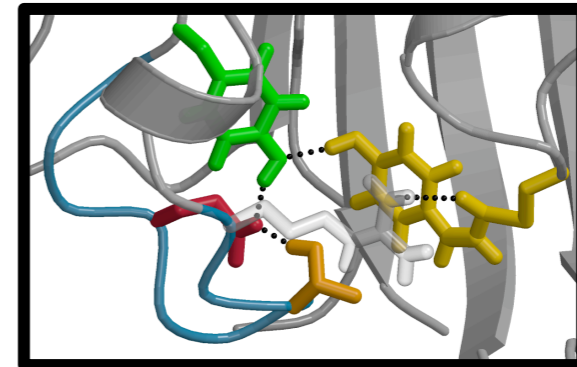
pB'



I $_{\alpha}$

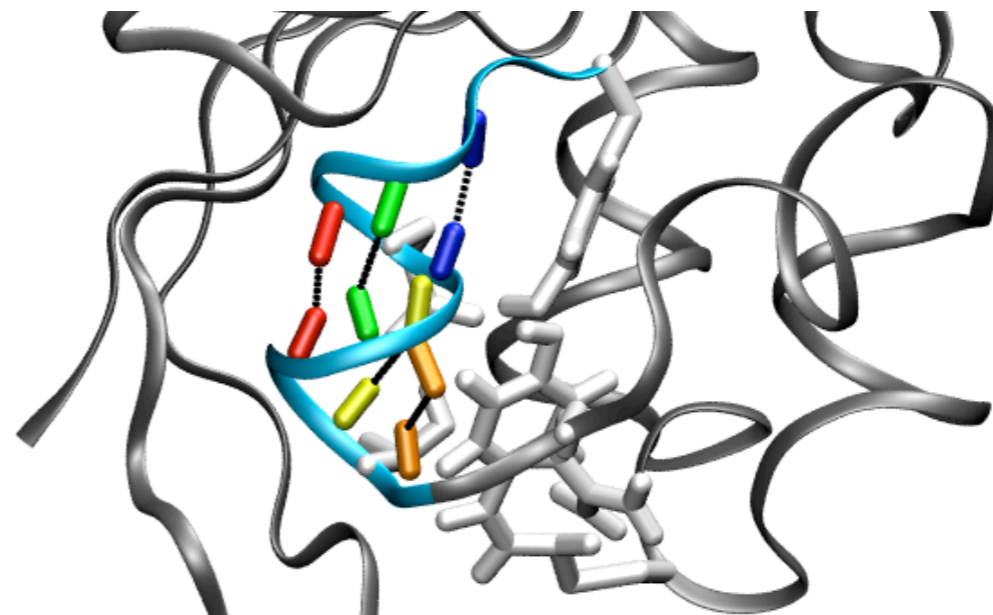


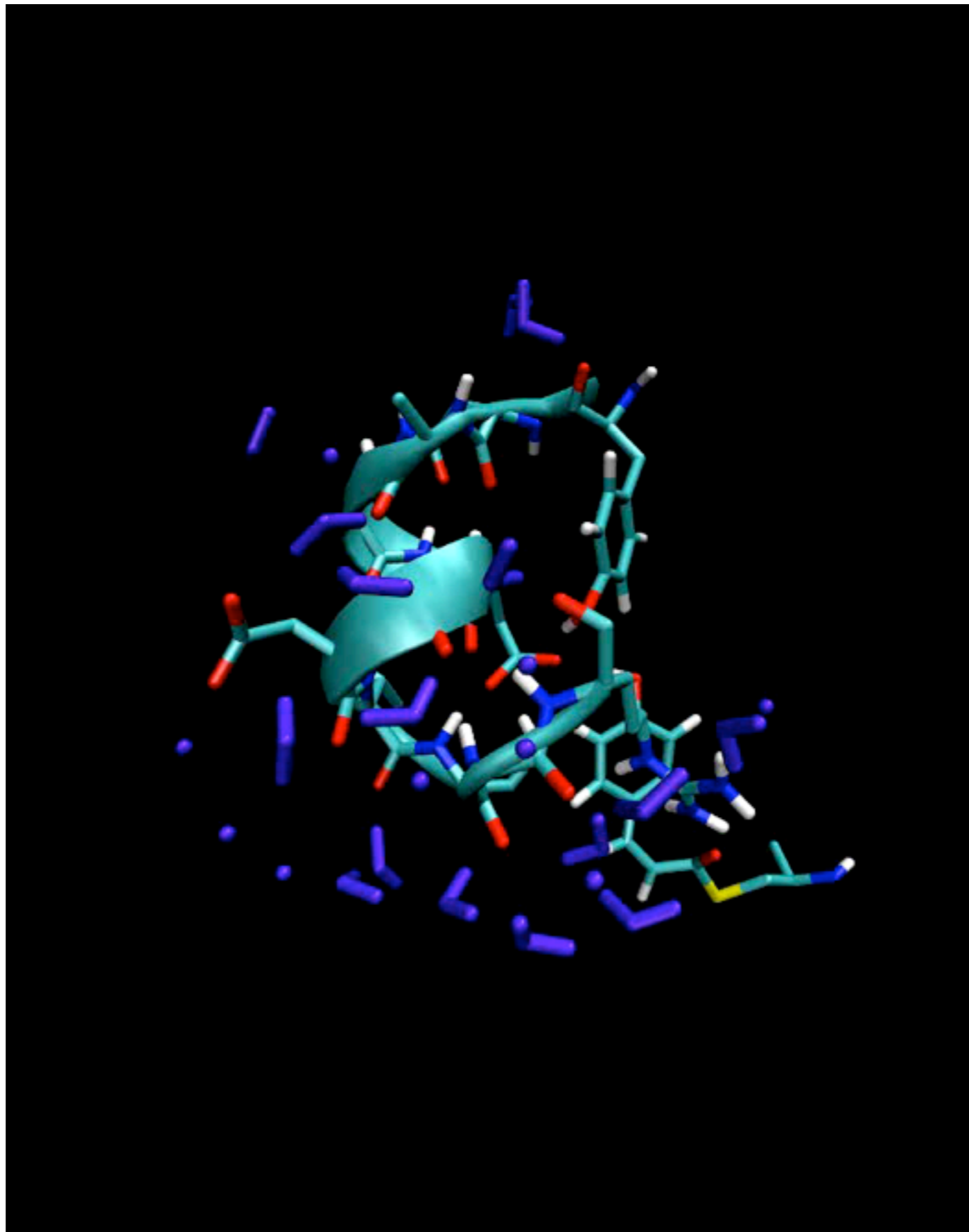
U $_{\alpha}$

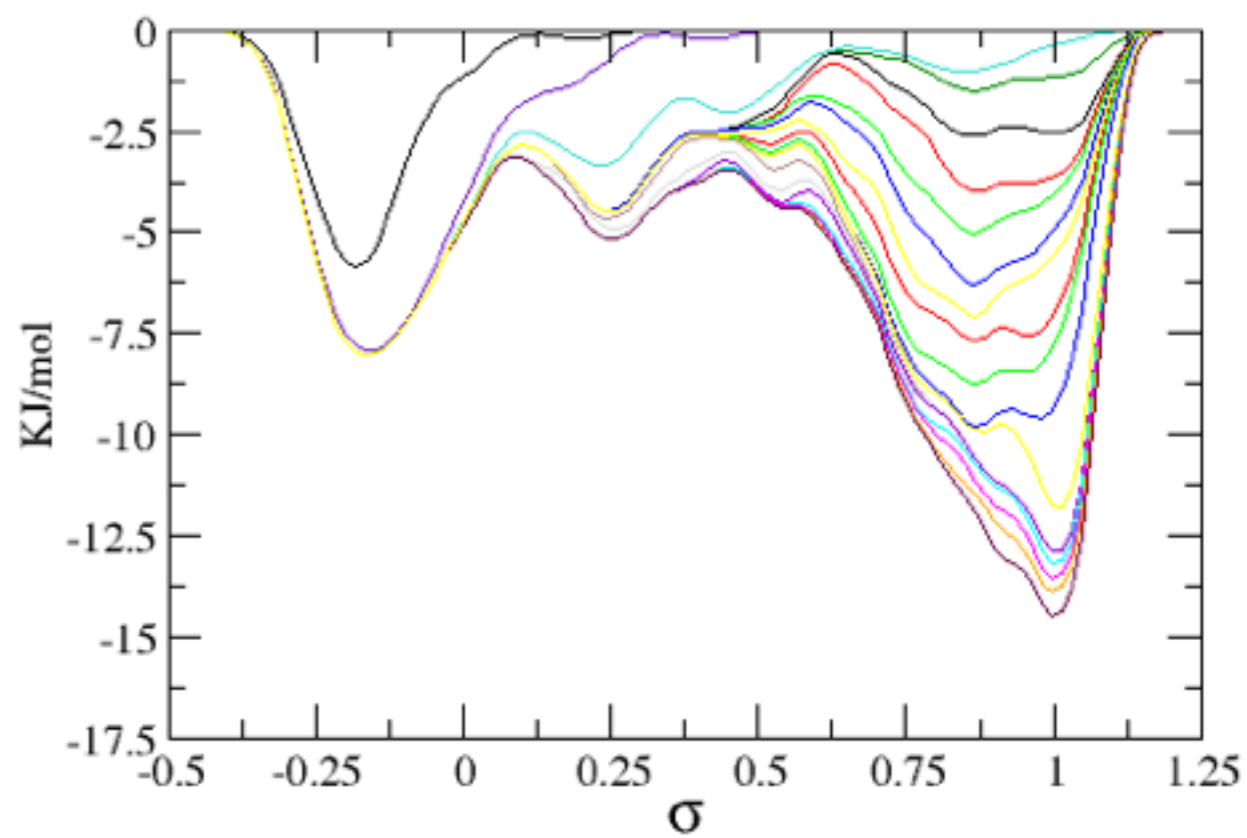
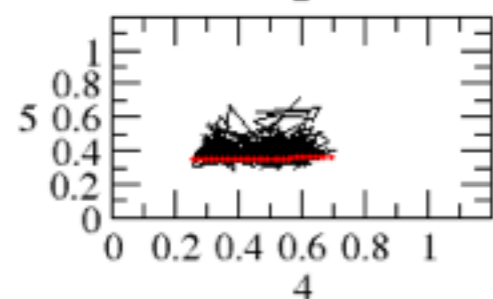
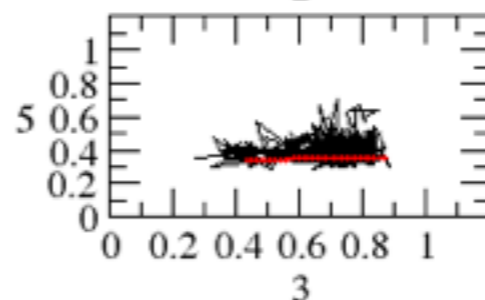
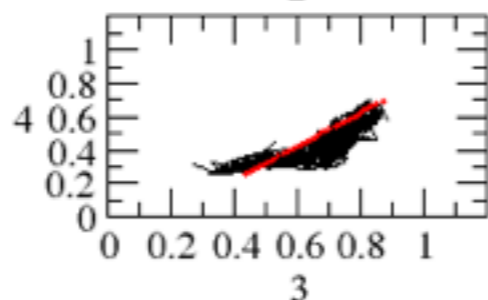
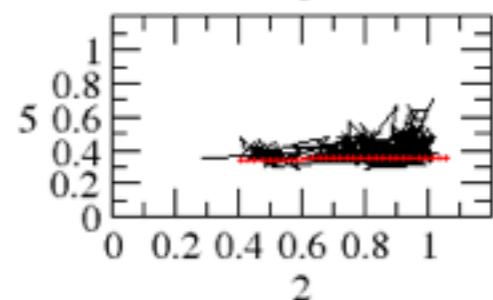
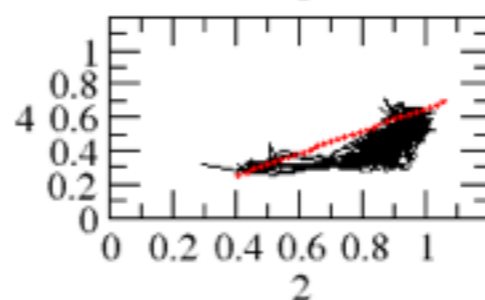
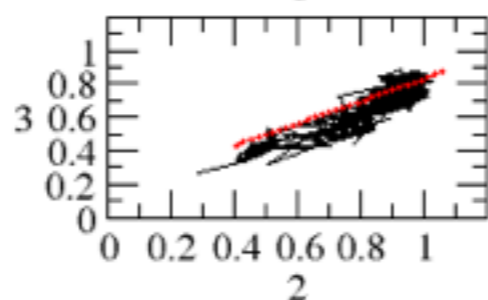
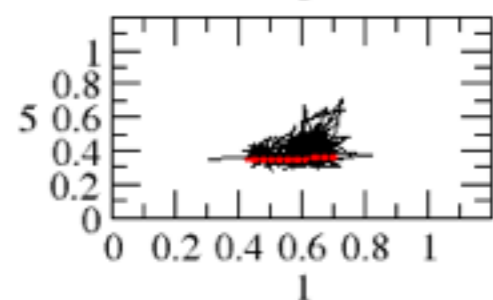
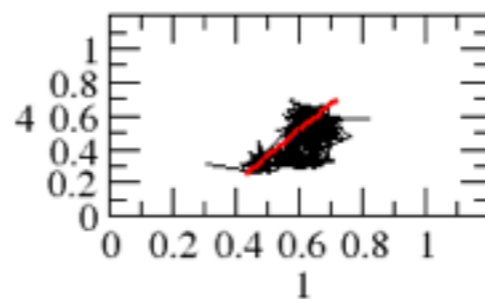
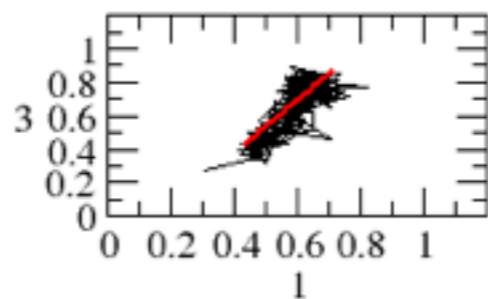
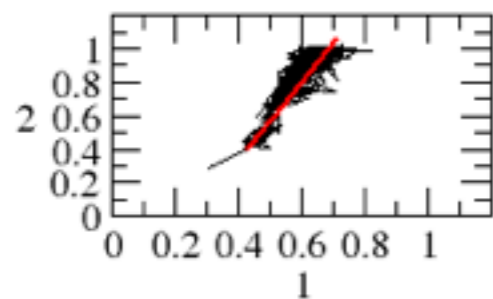


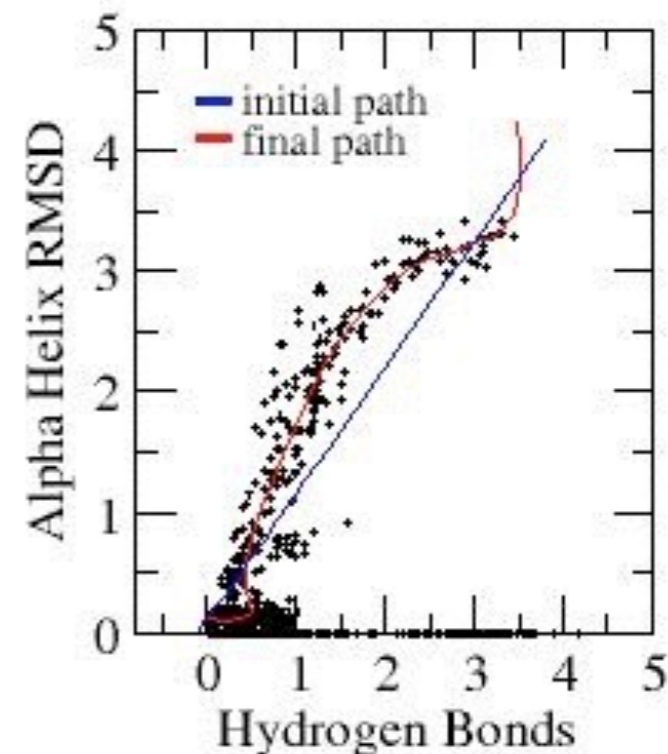
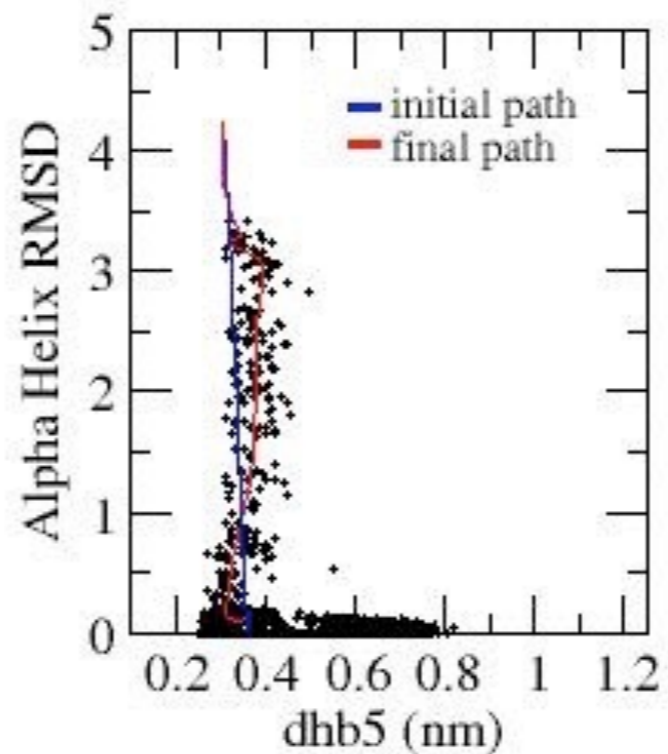
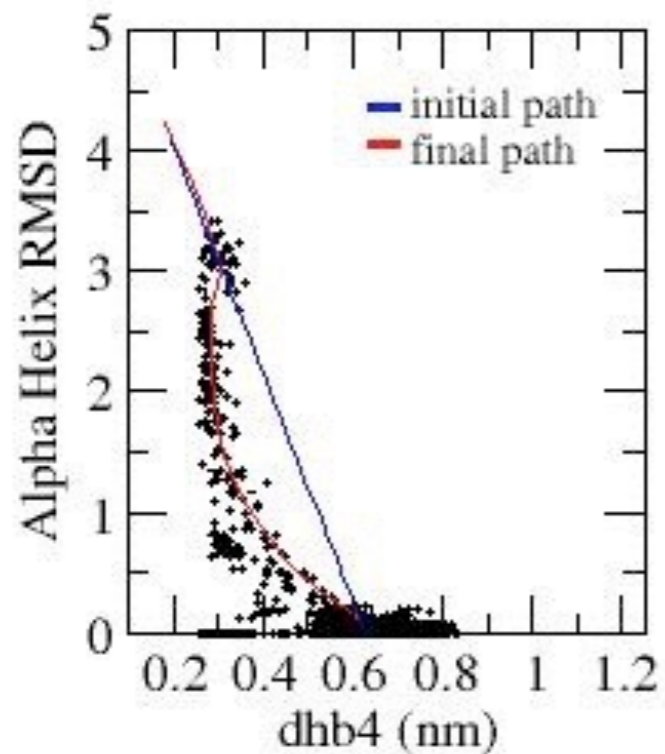
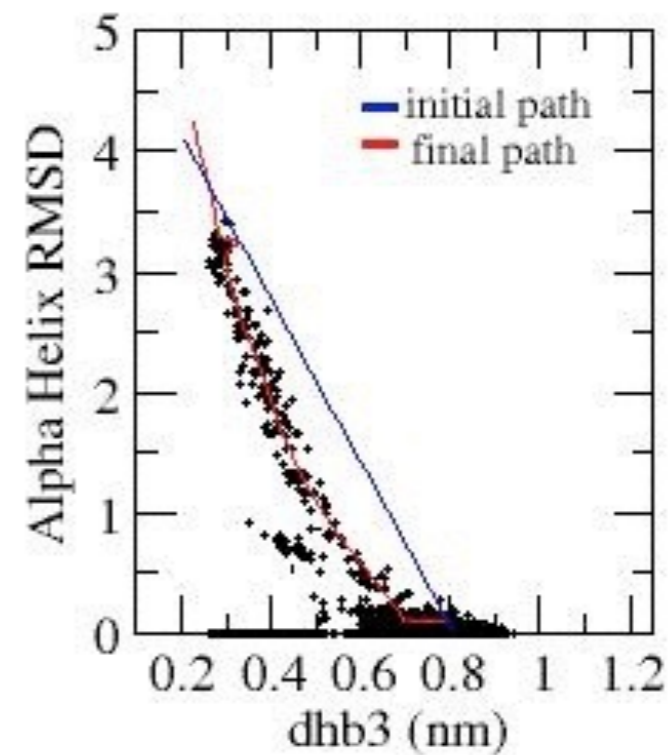
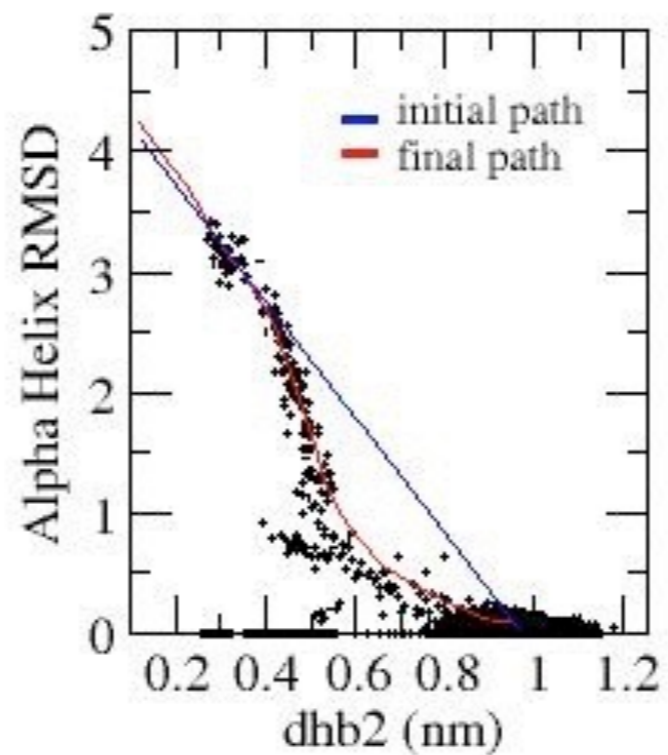
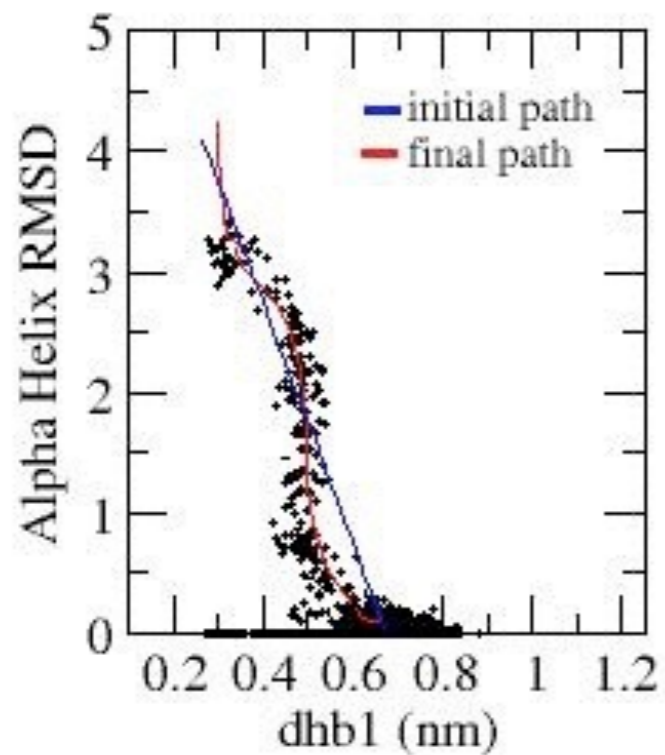
Collective variables:

- Asn43-O---Gly47-H = $\text{dhb}_{\alpha 3}^1$
- Ala44-O----Asp48-H = $\text{dhb}_{\alpha 3}^2$
- Ala45-O----Ile49-H = $\text{dhb}_{\alpha 3}^3$
- Glu46-O----Thr50-H = $\text{dhb}_{\alpha 3}^4$
- Gly47-O----Gly51-H = $\text{dhb}_{\alpha 3}^5$

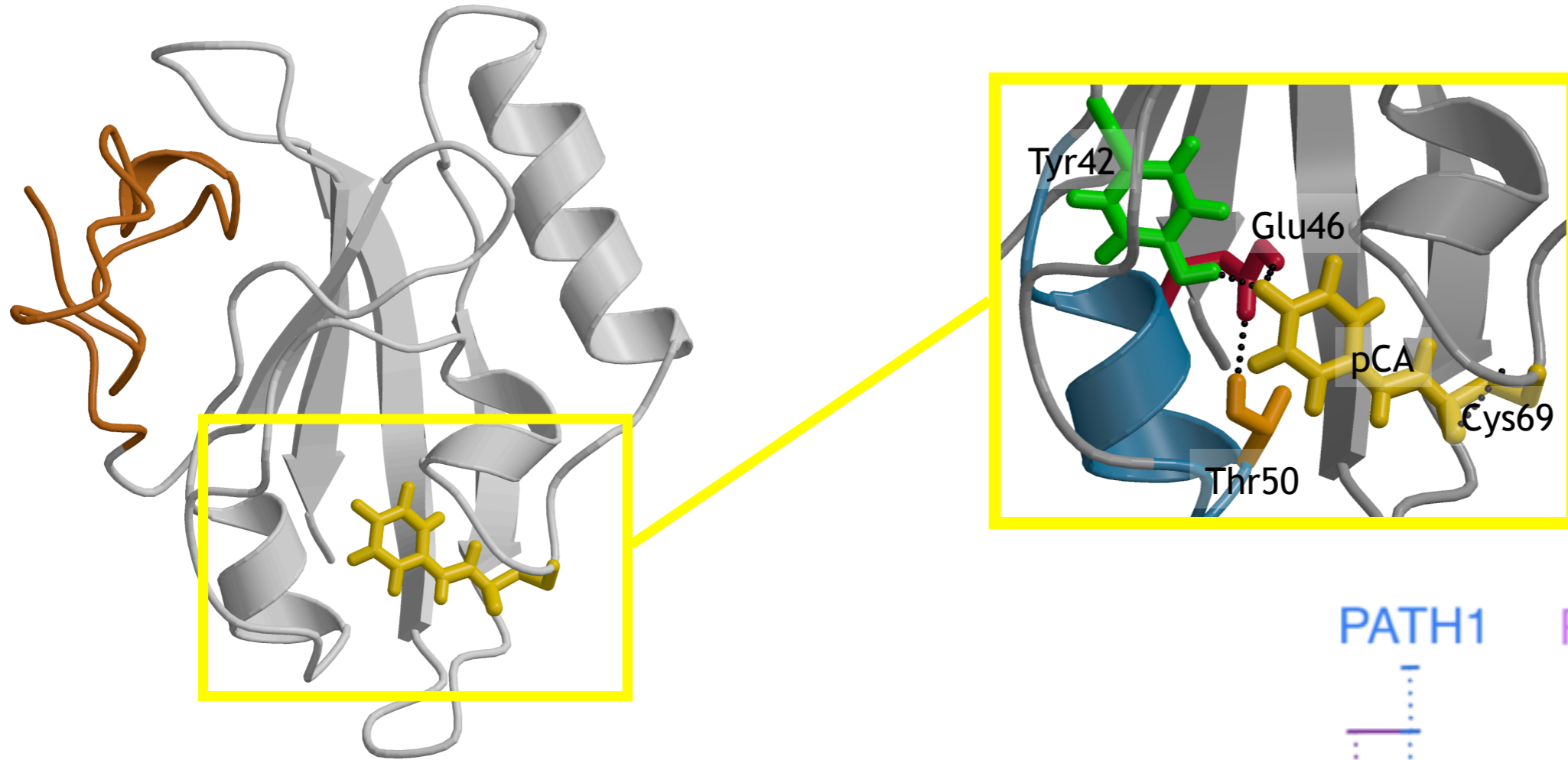








TPS of proton transfer process

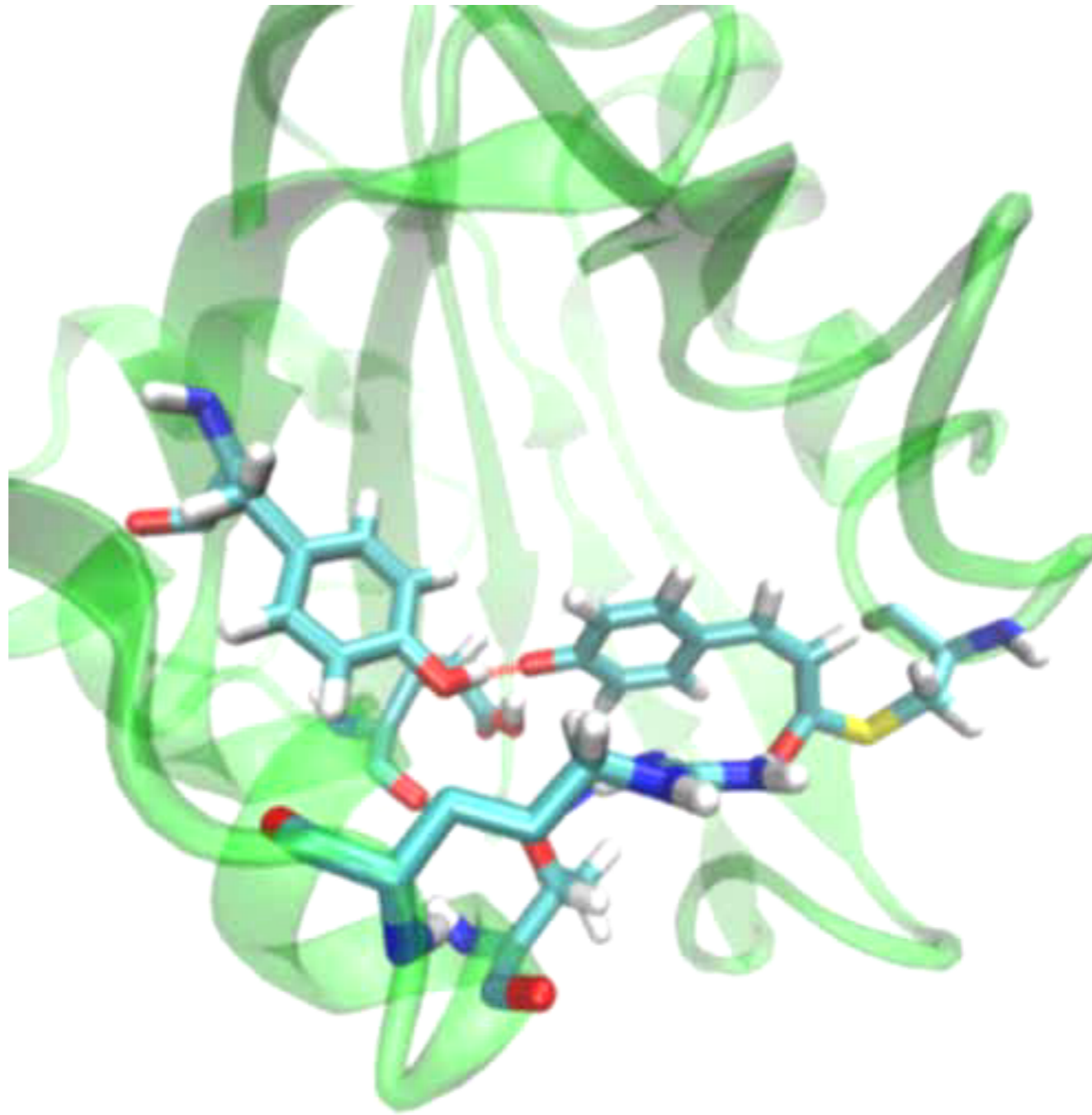


- CPMD QM/MM
- two-way shooting TPS
- 2 criteria:
 - Gln-H-CMA transfer
 - Tyr42 H-bond shift
- ~100 accepted paths
- variable path length

$$P_{acc}[\mathbf{x}^{(0)} \rightarrow \mathbf{x}^{(n)}] = h[\mathbf{x}^{(n)}; \mathbf{T}^{(n)}] \min \left[1, \frac{L^{(0)} \rho(\mathbf{x}_t^{(n)})}{L^{(n)} \rho(\mathbf{x}_t^{(0)})} \right]$$

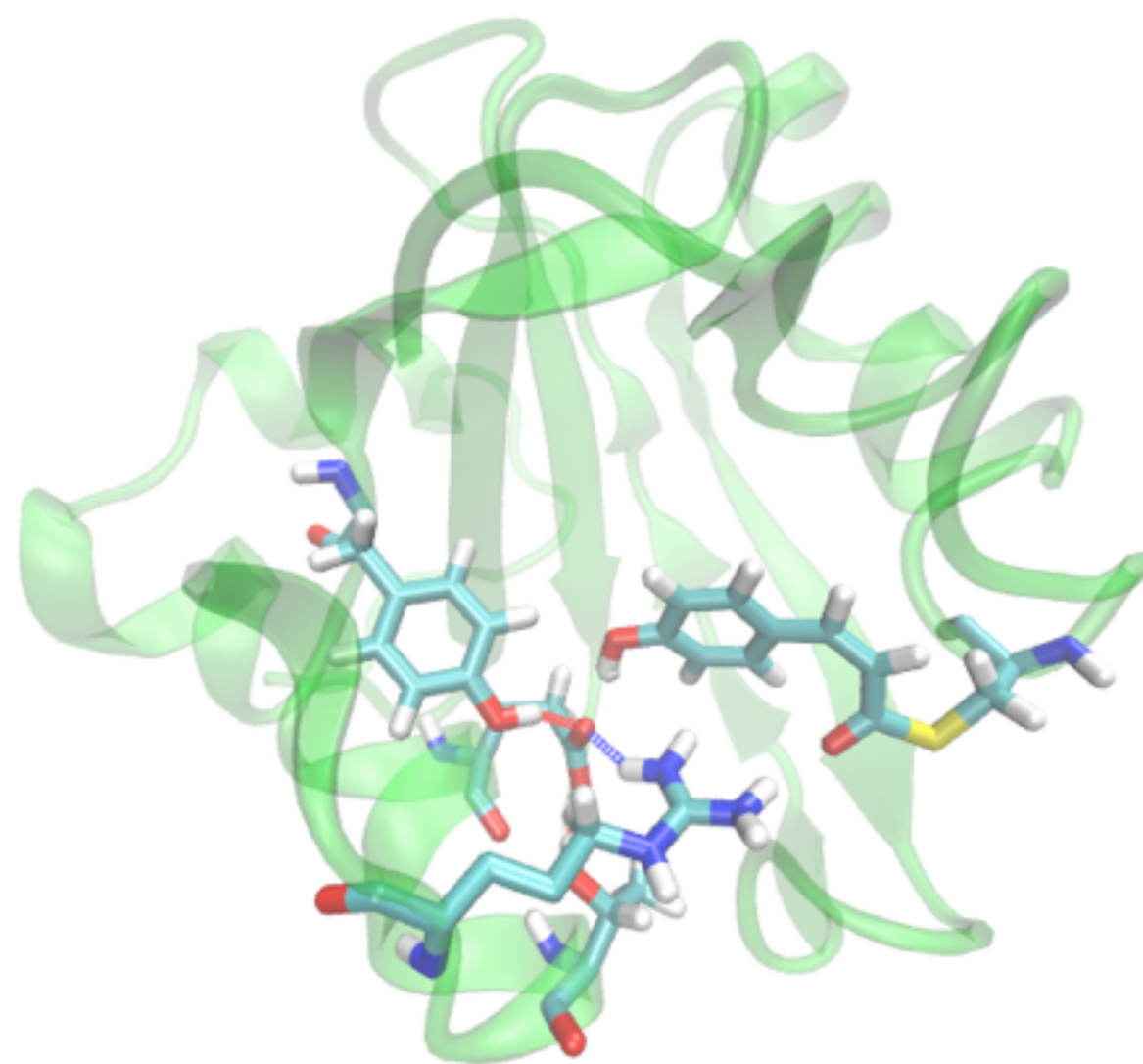
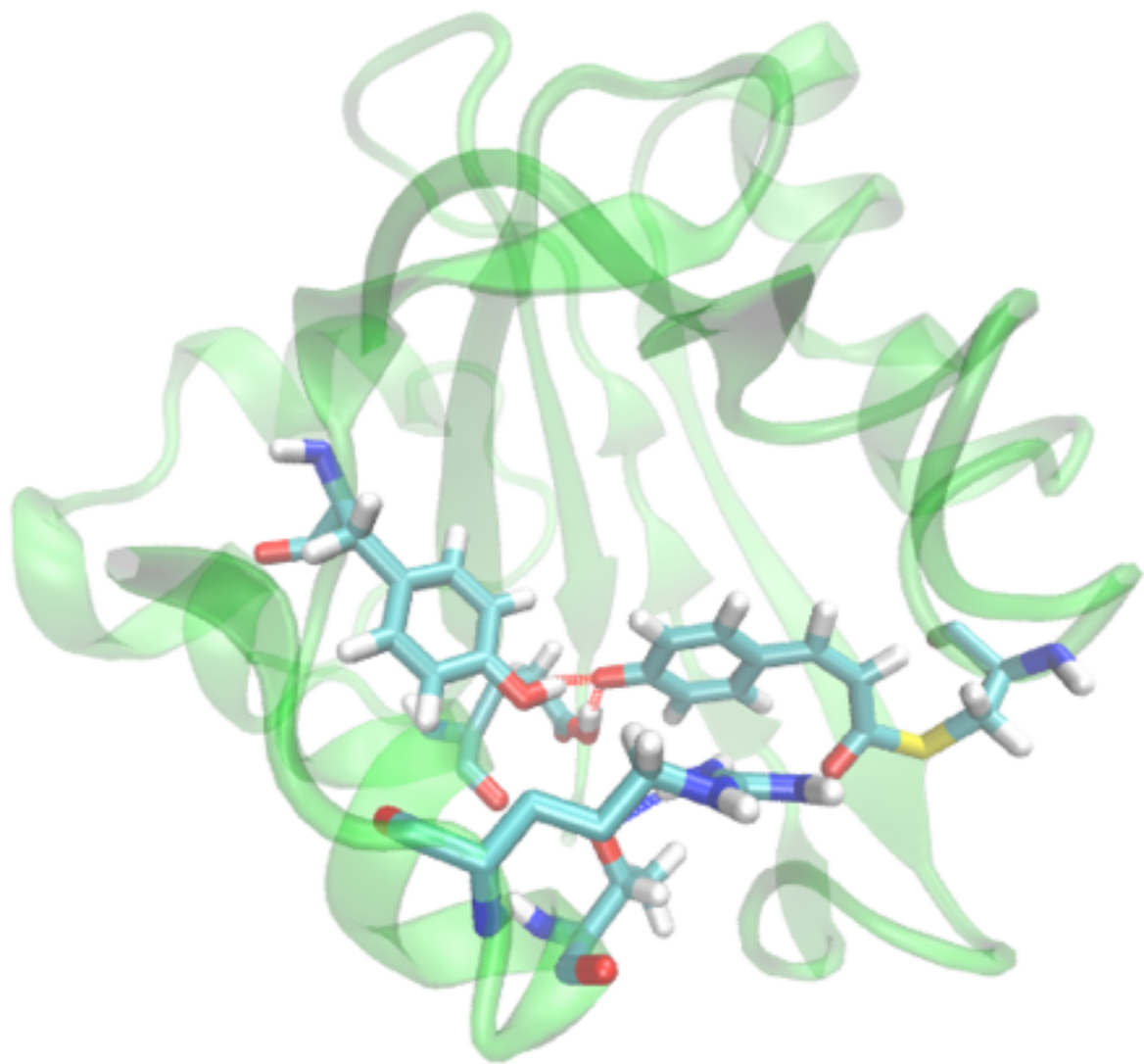


TPS of proton transfer process



TPS of proton transfer process

pR to pB'



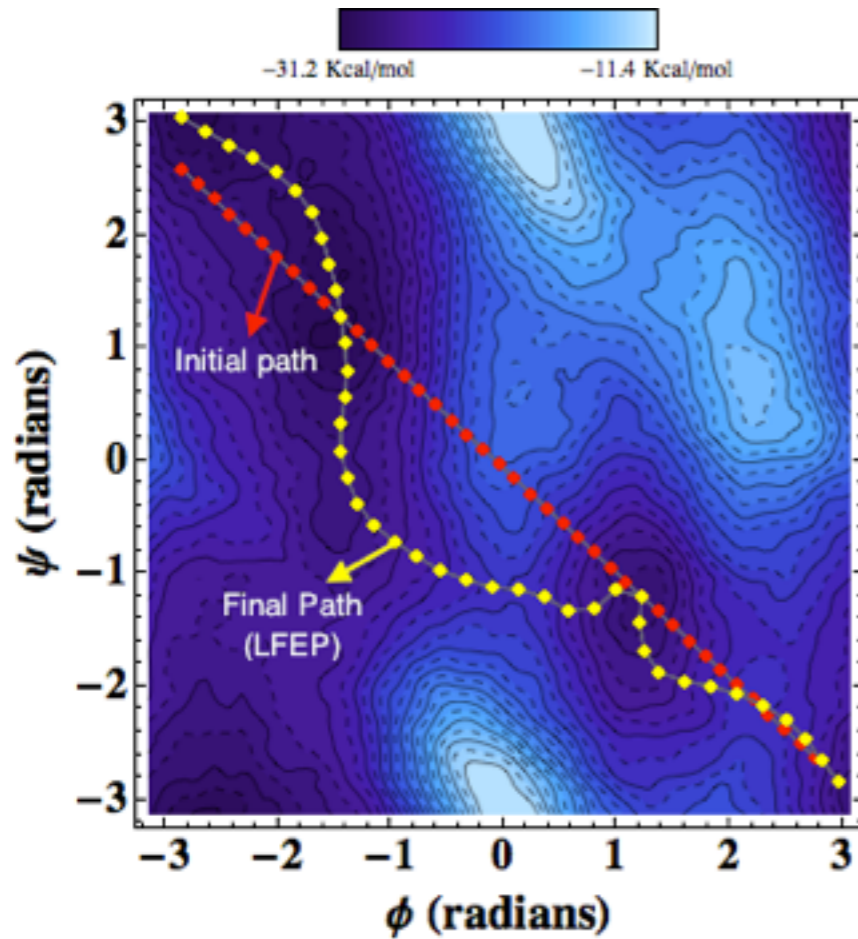
practical implementation

why it works

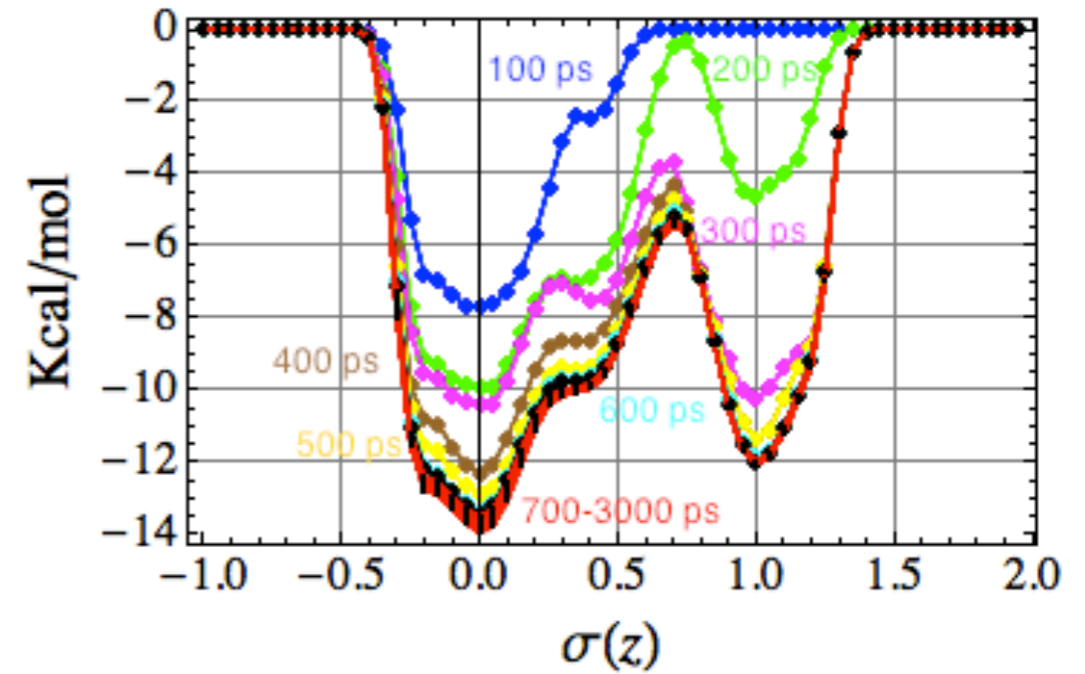
- small τ gives dynamic path nodes
- growing potential pushes along path
- self-healing bias-potential

path-metadynamics convergence

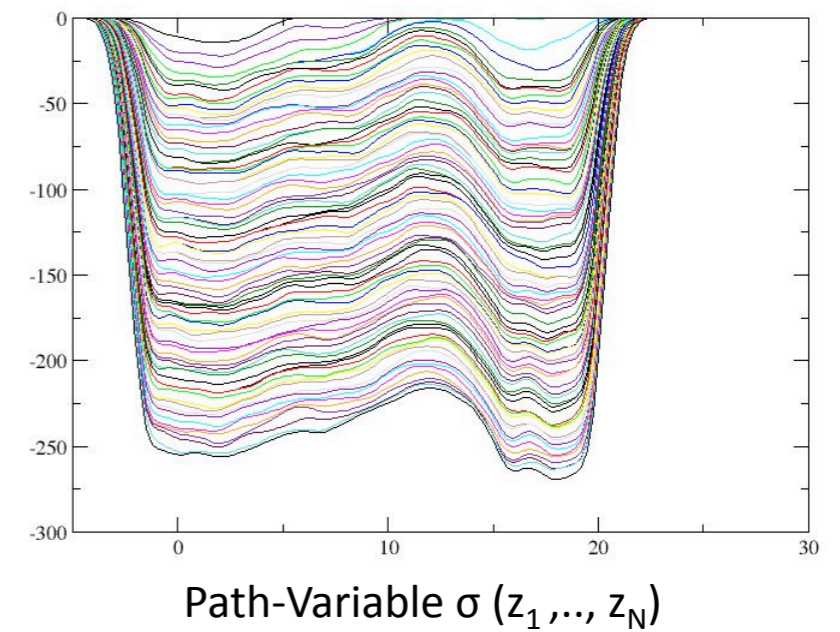
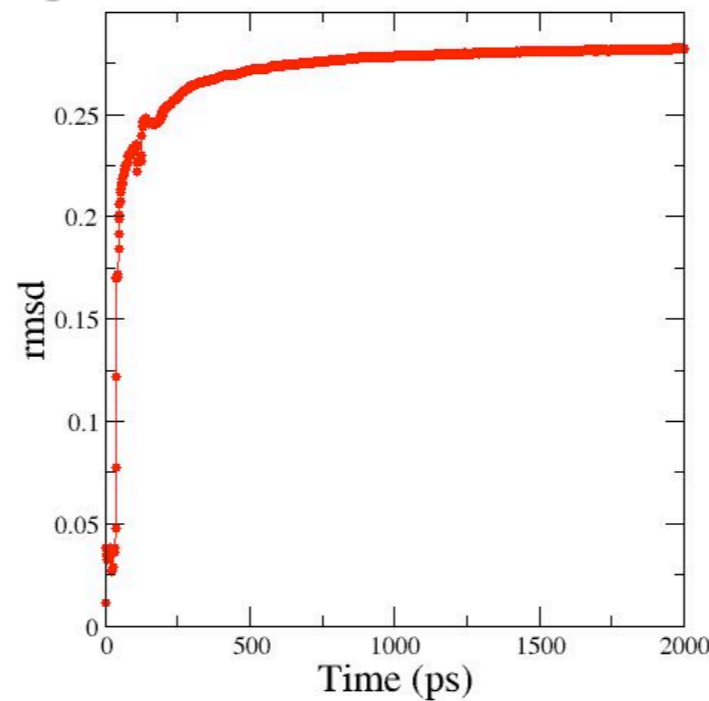
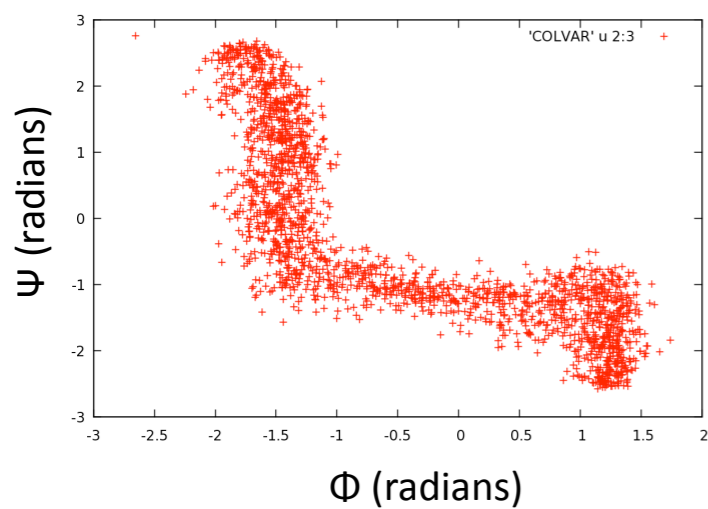
Monitoring convergence of the path



Free Energy along the path



$$d(t) = \frac{1}{R} \sqrt{\sum_{p=1}^R \|s^p(t) - s^p(0)\|}$$



- The Liouville Operator
- MD Integrators

- Rare event simulations
- Reactive flux method

- Free energy methods
- path-metadynamics